EU request for a Technical Service to produce a compilation of assessment methods and indicators that can be used to assess seabed habitats under D6/D1 for the MSFD

Service summary

This Technical Service is in response to an EU request to review benthic indicators and assessment methods used under the MSFD to assess the state/condition of seabed habitats. The review was conducted with input from the ICES Workshop to scope assessment methods to set thresholds and assess adverse effects on seabed habitats (WKBENTH2, 24-26 May and 8-10 June 2022). The final output of this Technical Service (Annex 1) is a compilation of the most operational assessment methods and indicators for seabed habitats under D6/D1 for the MSFD with all relevant information to compare the methods and indicators between one another.

Request

ICES is requested to provide a detailed review of indicators used, or under development, by Regional Sea Conventions (RSCs), Member States and ICES, for assessing the state/condition of seabed habitats and relevant existing literature. The review should focus on peer-reviewed indicators which have large-scale application, and should also consider methods that have been developed for assessing the state/condition of seabed habitats suitable for MSFD assessments. The review of methods should include indicators based on both direct observational data and on models. Relevant indicators to be reviewed include those of RSCs for quality status assessments, of Member States for MSFD purposes such as under the Water Framework Directive (WFD) and the Habitats Directive (HD), and those used by ICES. The review should specify the input data, how it is processed, the parameters of habitat quality used, how quality is quantified, any threshold values used, the applicable seabed (habitat) and pressure types, how the output is expressed, and how confidence and uncertainty are handled.

Outputs

Over 600 assessment methods and indicators for seabed habitats under D6/D1 for the MSFD were considered from various sources, including EU-funded projects, Regional Seas Conventions, and scientific literature. With input from WKBENTH2, this technical service consists of the most operational assessment methods and indicators (Table 1) to assess seabed habitats under D6/D1 for the MSFD. Supporting information for each of the identified assessment methods and indicators (Table 1) is compiled in a standardized template, contained in Annex 1. This information will facilitate the comparison of methods and indicators between one another, and will help evaluating assessment methods to set thresholds and assess adverse effects on seabed habitats under D6/D1 for the MSFD.

Table 1. Overview of benthic indicators or assessment methods. The indicators and methods stated in italics are not (yet) included in the indicator compilation presented in the Annex, as input from experts is still underway.

User	Benthic indicator or assessment method
OSPAR	BH1 – Sentinels of the Seabed (SoS)
	BH2b – Margalef diversity index
	BH3 – Extent of physical disturbance to benthic habitats
	BH4 – Area of Habitat Loss
	BISI – Benthic Indicator Species Index
ICES	L1 – Fraction community longevity exceeding trawling interval
	L2 – Reduction in median community longevity
	PD – Population Dynamics Model
HELCOM	CumI – Cumulative Impacts on benthic biotopes
	BQI – Benthic Quality Index, State of the soft-bottom macrofauna community
	Condition of benthic habitats
SPA/RAC	*EO1-CI1 – Ecosystem Objective 1: Biodiversity, Common Indicator 1*
	EO1-CI2 – Ecosystem Objective 1: Biodiversity, Common Indicator 2

User	Benthic indicator or assessment method	
Member States	Belgium / Netherlands	BEQI – Benthic Ecosystem Quality Index
	Denmark	DKI – Danish Quality Index
	Estonia	ZKI – Estonian coastal water macrozoobenthos community index
	Finland	BBI – Brackish water Benthic Index
	France	TDI – Trawling Disturbance Index
	France	mTDI – Modified Trawling Disturbance Index
	France	pTDI – Partial Trawling Disturbance Index
	France	mT – Modified Vulnerability Index
	Germany	MarBIT – Marine Biotic Index Tool
	Greece	BENTIX
	Ireland / UK	IQI – Infaunal Quality Index
	Norway	*NQI – Norwegian Quality Index*
	Portugal	*BAT – Benthic Assessment Tool*
	Spain	*BOPA – Benthic Opportunistic Polychaetes Amphipods Index*
	Spain	MEDOCC – MEDiterranean OCCidental
General	GPBI – General Purpose Benthic Index	
	AMBI – AZTI's Marine Biotic Index	
	M-AMBI – Multivariate AZTI's Marine Biotic Index	
	NEAT-tool – Nested Environmental status Assessment Tool	

Basis of the technical service

This technical service is established in close collaboration with the WKBENTH2 workshop, held in May (24-26) and June (8-10) 2022. During this workshop, both the framework for compiling relevant information on benthic indicators and assessment methods, and the final list of selected benthic indicators and assessment methods was agreed upon.

Sources and references

ICES. 2022. Workshop to scope assessment methods to set thresholds (WKBENTH2). ICES Scientific Reports. 4:70. 99 pp. http://doi.org/10.17895/ices.pub.20731537

Recommended citation: ICES. 2022. EU request for a Technical Service to produce a compilation of assessment methods and indicators that can be used to assess seabed habitats under D6/D1 for the MSFD. *In* Report of the ICES Advisory Committee, 2022. ICES Advice 2022, sr. 2022.11. https://doi.org/10.17895/ices.advice.21070975.

Annex 1. Compilation of the most operational assessment methods and indicators.

OSPAR: BH1 / SoS

Indicator description	Indicator name	SoS (BH1)			
	Indicator description	Sentinel of the Seabed (SoS) indicator, called BH1 in OSPAR framework. The indicator assesses the status of benthic indicator by computing the proportion of sentinel species across a pressure gradient (pressure-state curves). Sentinel species are defined by the method as species which are sensitive to the pressure and typical of the habitat under reference conditions.			
	Type of indicator	X Model	☐ Empirical-based	☐ Pressure	
	Pressure assessed	Physical abrasion			
	Human activity	Demersal fisheries / bottom trawling			
	MSFD criteria /descriptor	D6 (D6C3 and D6C5)			
	How does the indicator relate to benthic biological diversity?	The indicator measured the state of benthic habitats based on the proportion of sensitive species. Low values of the indicator compared to reference conditions means reduction in the proportion of these species and therefore potentially diversity lost.			
	How does the indicator relate to benthic community structure and function?	When measuring trawling impacts the SoS indicator uses the BESITO index to rank species according to its sensitivity to the pressure. This indicator uses biological traits to establish this sensitivity. Structural habitat forming species are usually ranked as highly sensitive and quite often selected as sentinel species (see Serrano et al., In press). Furthermore, because is based on biological traits, to measure functional lost (e.g. reduction in proportion of filter species or reduction in burrower benthos) can be easily computed within the frame of the indicator although is not a direct output.			
	Indicator status	X Under development	☐ Applied for MSFD	☐ Applied for other management, if so, for what:	
	Regions with operational assessments	The indicator has been tested across several areas of the Atlantic and Mediterranean (Serrano et al., In Press) and is currently been used in region IV in OSPAR as well as in the Spanish Mediterranean waters.			
Input data	Biological data input (e.g. monitoring program, time series, sampling method)	Benthic species information for monitoring data. It has been tested with different types of methods (video surveys, box-corer and trawling samples). When using to measure trawling impacts It also need biological traits information.			
	Targeted organisms	x Infauna	x Epifauna	☐ Demersal fish	☐ Other: ….
	Environmental data input (e.g. empirical/modelled, source, time series)	MSFD broad habitat layer			
	Pressure data input (e.g. time series, empirical/modelled, source, national/international)	Swept area ratio per year			
	Data availability	The indicator is an OSPAR common indicator in region IV. Preliminary assessment outputs already available in OSPAR website. Final assessment of region IV planed for the end of this year.			
Methodology	Parameters determined from biological data (e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)	Proportion of sentinel species by grid cell			
	Parameters determined from pressure data (e.g. total SAR, years not fished, trawling interval)	Annual average (surface) Swept Area Ratio per grid cell			
	Algorithm type List of categorical information (Presence/Absence, …) Direct measurements (counts, areas, concentrations, …) Single or multimetric indicators using basic arithmetics Indicators using multivariate and complex statistics	3 Single basic metric (proportion of sentinel species)			

	Indicators derived from modelling approaches Indicators reporting on trends	
	References for state-pressure relation	Serrano et al (Accepted in Ecological indicators, https://doi.org/10.1016/j.ecolind.2022.108979) BH1 CEMP Guidelines (already provided)
	Uncertainty estimation methodology	Yes, the standard error associated to the pressure.
	Coding availability (e.g. scripts, GitHub)	Code available at: https://github.com/Gonzalez-Irusta/SoS
	Threshold present	Yes
	Threshold methodology	The methodology is under development and is not yet considered final. In our approach we are exploring the use of the pressure-state curves generated based on the correlation computed (using GAMs) between the proportion of sentinel species and the pressure to define two things: i) Habitat sensitivity (based on the relationship between the curved obtained with the GAMs and five theoretical curves for 5 different sensitivities, see Serrano et al., 2022) and ii) the tipping point, as a proxy to the point under which the habitat can not lost more quality (significantly). Finally, we will suggest different quartiles, between cero and the tipping point depending of habitat sensitivity, being more close for habitat less sensitive and viceversa. These values (quartiles) are still arbitrary so further information to do this differently and more scientifically robust will be very welcome.
Output	Output variable type	x Continuous ☐ Categorical ☐ Proportional
Output	Output variable range / classes	0 (community lost/destroyed) to 1 (unimpacted)
Output	Output availability (e.g. report, website, reference)	The indicator is an OSPAR common indicator in region IV. Preliminary assessment outputs already available in OSPAR website. Final assessment of region IV planed for the end of this year.
Output	Uncertainty handling (e.g. present confidence interval)	Associated standard error map to the SoS prediction is provided
Output	Spatial resolution (e.g. grid cell size, habitat level)	0.05 x 0.05 latitude x longitude degree
Output	Temporal resolution	Once; determining annual average over the used time period
Output	Seabed habitat levels presented?	MSFD broad habitat types
More info	Indicator lead person	Alberto Serrano a.serrano@csic.ieo.es José Manuel González Irusta jmanuel.gonzalez@csic.ieo.es
More info	Indicator data contact	José Manuel González Irusta jmanuel.gonzalez@csic.ieo.es
More info	References / Literature / Project websites	OSPAR CEMP Guidelines González-Irusta, J. M., De la Torriente, A., Punzón, A., Blanco, M., & Serrano, A. (2018). Determining and mapping species sensitivity to trawling impacts: the BEnthos Sensitivity Index to Trawling Operations (BESITO). ICES Journal of Marine Science, 75(5), 1710-1721. Serrano, A., de la Torriente, A., Punzón, A., Blanco, M., Bellas, J., Durán-Muñoz, P., ... & González-Irusta, J. M. (2022). Sentinels of Seabed (SoS) indicator: Assessing benthic habitats condition using typical and sensitive species. https://doi.org/10.1016/j.ecolind.2022.108979

OSPAR: BH2b

Indicator description	Indicator name	Margalef diversity (BH2b)		
Indicator description	Indicator description	Margalef's index of diversity calculated as species richness (S-1) divided by abundance (ln(N)); a relative value (Relative Margalef diversity) is calculated by dividing the Assessed Margalef by a Reference Margalef (defined at the level of Assessment Units (AU), Broad Habitat Types (BHT), Monitoring techniques, etc, or combinations of those) to make results comparable. ('Relative Margalef diversity' used to be called 'Normalized Margalef diversity' before).		
Indicator description	Type of indicator	☐ Model	X Empirical-based	☐ Pressure
Indicator description	Pressure assessed	Basically, the general quality status (result of all pressures that are at stake) is assessed. Margalef diversity has been used with regards to fishing pressure (physical disturbance) and eutrophication (organic enrichment) but can potentially be used in any pressure gradient if sufficient data to identify		

		reference values are present. (Once reference values have been identified before for certain cases (AU x BHT x technique), those can be reused and basically no information on pressure distributions is needed to run the assessments).			
	Human activity	Demersal fisheries/bottom trawling or organic enrichment or inorganic pollutants.			
	MSFD criteria /descriptor	D6 (D6C5)			
	How does the indicator relate to benthic biological diversity?	It is calculated benthic biological diversity calculated based on community observation data, taking reference diversity into account.			
	How does the indicator relate to benthic community structure and function?	Benthic diversity relative to the reference is in fact a measure of benthic community structure; a relation between benthic diversity and functioning is expected (although the indicator does not include information on specific valuable species with certain functions).			
	Indicator status	X Under development (At the moment the methodology is being updated for the QSR2023 to match with BHTs and there are slight adjustments to the calculation of the reference)	X Applied for MSFD		X Applied for other management, if so, for what: OSPAR IA2017 & QSR2023 (in progress)
	Regions with operational assessments	Southern North Sea (IA2017), now being extended to the entire (Greater) North Sea region (QSR2023). (Potentially applicable to any region in case sufficient data from low pressure areas are available).			
Input data	Biological data input (e.g. monitoring program, time series, sampling method)	Benthic community data of any kind; has been applied on grab and core data; at the moment also being applied on benthic dredge and trawl data (in case references still have to be identified, data from low pressure areas and/or preferably data from several years should be available).			
Input data	Targeted organisms	X Infauna	X Epifauna	☐ Demersal fish	☐ Other: ….
Input data	Environmental data input (e.g. empirical/modelled, source, time series)	Variable – typically habitat maps are used (e.g. BHTs), but also areas distinguished on certain environmental conditions determining type of communities or species distributions could be used.			
Input data	Pressure data input (e.g. time series, empirical/modelled, source, national/international)	In case no reference values for Margalef diversity for the distinguished habitats, units and techniques have been defined before; information on pressure distribution is needed to identify low pressure areas. This can be SAR data or nutrient/pollutant levels, dependent of what is the dominant pressure.			
Input data	Data availability	IA2017: https://odims.ospar.org/documents/189/download (Benthic community data and data selected for reference sets) For backgrounds see: https://oap.ospar.org/en/ospar-assessments/intermediate-assessment-2017/biodiversity-status/habitats/condition-of-benthic-habitat-defining-communities/subtidal-habitats-southern-north-sea/, and https://doi.org/10.1016/j.ecolind.2017.09.029 (including assessment results available for download).			
Methodology	Parameters determined from biological data (e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)	Relative Margalef diversity (species or alternative taxa diversity) per sample Reference diversity per Assessment Unit (Used to be presented as average Relative Margalef diversity per Assessment Unit; currently median Relative Margalef diversity per Assessment Unit x Broad Habitat type (AU x BHT) is preferred).			
Methodology	Parameters determined from pressure data	Pressure input layers are retrieved from other sources (e.g. average subsurface fishing activity (subSAR per c-square) used for IA2017; average SAR per c-square in use for QSR2023 (https://doi.org/10.17895/ices.data.8192)).			

	(e.g. total SAR, years not fished, trawling interval)	
	Algorithm type List of categorical information (Presence/Absence, …) Direct measurements (counts, areas, concentrations, …) Single or multimetric indicators using basic arithmetics Indicators using multivariate and complex statistics Indicators derived from modelling approaches Indicators reporting on trends	3 (In case low pressure data for certain assessment units or habitats are scarce, reference values are retrieved using regression models; e.g. Reference Margalef - Median depth).
	References for state-pressure relation	Van Loon et al., 2018; https://doi.org/10.1016/j.ecolind.2017.09.029
	Uncertainty estimation methodology	Regression modelling of Margalef diversity to pressure levels (providing pseudo-R2 and significance levels). Standardized qualification of level of representativity of input data (number of samples, years, disturbance level in case of identification of reference value).
	Coding availability (e.g. scripts, GitHub)	https://cran.r-project.org/web/packages/BENMMI/index.html
	Threshold present	No
	Threshold methodology	-
Output	Output variable type	X Continuous ☐ Categorical ☐ Proportional
Output	Output variable range / classes	0 (no diversity; community lost or only one species present) to 1 (unimpacted; community at reference level) (In practice Relative Margalef diversity can locally be >1 (above reference level) in which case value is adjusted to ‘1’ (= reference state).
Output	Output availability (e.g. report, website, reference)	https://oap.ospar.org/en/ospar-assessments/intermediate-assessment-2017/biodiversity-status/habitats/condition-of-benthic-habitat-defining-communities/subtidal-habitats-southern-north-sea/ and https://doi.org/10.1016/j.ecolind.2017.09.029
Output	Uncertainty handling (e.g. present confidence interval)	Assessment results presented with 90% confidence limits
Output	Spatial resolution (e.g. grid cell size, habitat level)	Basically, assessment results per sample; for overall assessment purposes results are presented as average values at the level of Assessment Units and/or Broad Habitat Types as well.
Output	Temporal resolution	Basically, a result is obtained for an assessment period (e.g. 6-year MSFD period or OSPAR review period), but a series of assessment periods provides information on benthic habitat quality (as indicated by Relative Margalef diversity) development (e.g. currently result for the series QSR2010, IA2017, QSE2023 provide information on quality status development).
Output	Seabed habitat levels presented?	At IA2017 assessments were done at the level of Assessment Units expected to be representative for (aggregations of) Broad habitat types (largely characterized by one of the BHTs). At present (for QSR2023), assessments are done at the MSFD BHT level.
More	Indicator lead person	Sander Wijnhoven; sander.wijnhoven@ecoauthor.net
More	Indicator data	Sander Wijnhoven; sander.wijnhoven@ecoauthor.net

	contact	
	References / Literature / Project websites	OSPAR IA2017: https://oap.ospar.org/en/ospar-assessments/intermediate-assessment-2017/biodiversity-status/habitats/condition-of-benthic-habitat-defining-communities/subtidal-habitats-southern-north-sea/ Van Loon et al. (2018): https://doi.org/10.1016/j.ecolind.2017.09.029 BENMMI user manual: https://puc.overheid.nl/PUC/Handlers/DownloadDocument.ashx?identifier=PUC_151859_31&versienummer=1 Case study Denmark (Van Loon et al., 2018): https://puc.overheid.nl/PUC/Handlers/DownloadDocument.ashx?identifier=PUC_158511_31&versienummer=1

OSPAR: BH3

Indicator description	Indicator name	BH3 – Extent of physical disturbance to benthic habitats		
	Indicator description	Physical disturbance can damage seafloor habitats, particularly those that support larger and more fragile species and/or those that take a long duration to recover. The Physical Disturbance to Benthic Habitats indicator, OSPAR Common Indicator: BH3, can be used to enable large-scale assessments and improve understanding of anthropogenic pressures in marine environments. The indicator has direct application to OSPAR and assessments of Good Environmental Status under the MSFD and the UK Marine Strategy, alongside other national reporting mechanisms across the Northeast Atlantic. The BH3 indicator assesses the spatial extent and magnitude of potential physical disturbance to benthic habitats caused by human activities, where a known pressure-activity link is established. The indicator combines pressure data with information on receptor sensitivity, derived from traits-based assessments of biological communities that characterize assessed biotopes. Both habitat and species sensitivity are considered within BH3, in terms of resistance (the ability to withstand a given pressure) and resilience (the ability to return to an unimpacted state), in response to assessed physical pressures. Outputs are expressed in spatial formats, with maps and GIS layers produced for sensitivity, pressure, and disturbance. In addition, summary statistics, such as the proportion of habitats under varying levels of disturbance are calculated and presented in tabular and graphical formats. The indicator could be applied at a variety of spatial scales and includes a calculation of confidence levels based on the quantity and quality of the data layers		
	Type of indicator	⌖ Model	☐ Empirical-based	☐ Pressure
	Pressure assessed	Physical disturbance		
	Human activity	Bottom-contacting fishing & commercial aggregate extraction		
	MSFD criteria /descriptor	D1 - Biological Diversity; D6 - Seafloor Integrity; D6.C3 (extent);		
	How does the indicator relate to benthic biological diversity?	This indicator relates indirectly to biodiversity, e.g., through assessments of potential disturbance to biodiversity components (benthic communities & habitats, derived from species-resolution sensitivity), rather than direct assessments via diversity indices (e.g., Simpson's Diversity).		
	How does the indicator relate to benthic community structure and function?	BH3 considers a range of input data associated with benthic communities (including structure & function) when assessing potential physical disturbance. Please see examples below of the key areas where benthic community data are applied. Biotope Extent & Habitat Distribution Habitat maps used to spatially analyze sensitivity are classified using records and known distributions of species and biotopes (e.g., using biological characteristics). Biotopes are defined as a combination of an abiotic habitat and its associated community of species and are therefore, directly informed by community structure and associated functions. Sensitivity BH3 assesses the sensitivity of key structural, functional, and characterizing species of benthic habitats in relation to a defined intensity of a given pressure. Sensitivity is assessed in terms of resistance (ability to withstand pressure) and resilience (ability to return to an unimpacted state, using key biological data, such as biological traits (e.g., life history and the ecology of the key and characterizing species). Please see below an example of the information used when assessing habitat sensitivity: **Category** **Description**		

<table>
<tr><td rowspan="3"></td><td></td><td colspan="4"><table><tr><td>Key structural species</td><td>The species provides a distinct habitat that supports an associated community. Loss/degradation of this species population would result in loss/degradation of the associated community.</td></tr><tr><td>Key functional species</td><td>Species that maintain community structure and function through interactions with other members of that community (for example through predation, or grazing). Loss/degradation of this species population would result in rapid, cascading changes in the community.</td></tr><tr><td>Important characteristic species</td><td>Species characteristic of the biotope (dominant, and frequent) and important for the classification of the habitat. Loss/degradation of these species populations may result in changes in habitat classification.</td></tr></table></td></tr>
<tr><td>Indicator status</td><td colspan="2">☐ Under development</td><td>⌖ Applied for MSFD: UK, DE, IE & references included by OSPAR Contracting Parties in MSFD assessments.</td><td>⌖ Applied for other management, if so, for what: OSPAR & UK Marine Strategy</td></tr>
<tr><td>Regions with operational assessments</td><td colspan="4">OSPAR Regions:
North Sea (II).
Celtic Seas (III).
Bay of Biscay and Iberian Coast (IV).
Wider Atlantic (V).

MSFD Regions:
Greater North Sea, including the Kattegat and the English Channel.
Celtic Seas.
Bay of Biscay and Iberian Coast.</td></tr>
<tr><td rowspan="4">Input data</td><td>Biological data input (e.g., monitoring program, time series, sampling method)</td><td colspan="4">Biological presence data (at species resolution) derived from benthic surveys & recurring monitoring schemes, including MPA monitoring across the OSPAR Area. Data are derived from data management systems, such as the JNCC Marine Recorder database and via data calls from across the OSPAR Area.

Biological data are sampled using a diversity of gears and approaches, dependent on the input/data provider. Examples can include, but are not limited to physical sampling, such as Hamon/Day grabs and benthic epifauna data obtained with trawls.

BH3 also considers biological traits information when assessing species and habitat sensitivity against a given pressure.</td></tr>
<tr><td>Targeted organisms</td><td>⌖ Infauna</td><td>⌖ Epifauna</td><td>☐ Demersal fish</td><td>☐ Other: ….</td></tr>
<tr><td>Environmental data input (e.g., empirical/modelled, source, time series)</td><td colspan="4">Environmental data are used within BH3 in the following formats:
Habitat maps created from survey data, obtained from EMODnet Seabed Habitats Data Portal.

Broad-scale predictive habitat maps comprising:
EUSeaMap 2021, which covers all European sea basins where the EMODnet Geology seabed substrate map is available. EUSeaMap 2021 also
UKSeaMap 2018, a revised version of EUSeaMap, which provides greater resolution of modelling accuracy for UK waters.
Data derived from EUSeaMap comprise a suite of EMODnet environmental datasets, including EMODnet Bathymetry, EMODnet Geology and Copernicus marine services via the Copernicus Marine Environment Monitoring Service (CMEMS).
Additional physical data used for the calculation of the models include data on light attenuation, light at the seabed and kinetic, current and wave energy datasets.</td></tr>
<tr><td>Pressure data input (e.g., time series, empirical/modelled, source, national/international)</td><td colspan="4">BH3 currently assesses potential physical disturbance from fishing and commercial aggregate extraction, although, the method can be adapted for wider activities known to cause physical pressure, where pressure and sensitivity data are available. Additional activities will be considered in forthcoming assessments.

Fishing pressure:
Total annual SAR values
Range SAR assessments, accounting for interannual SAR variability (please see BH3 CEMP Guidelines for further detail on fishing variability assessments). Range assessments are conducted across OSPAR QSR & MSFD intervals.</td></tr>
</table>

<table>
<tr><td></td><td></td><td>Areas not fished

Commercial aggregate extraction pressure:
Licensed extraction areas, with associated extraction statistics (where available).
Commercial aggregate dredging footprint and dredging intensity (gridded) as either:
The total volume dredged per licensed extraction area/per grid cell, or.
The total duration of extraction in units of time and volume dredged per grid cell as gridded spatial data, indicative of the activity intensity, including both vessel Automatic Identification System (AIS) and Electronic Monitoring System (EMS) data.
Confirmation of OSPAR Contracting Parties where aggregate extraction activity is known not to occur.</td></tr>
<tr><td></td><td>Data availability</td><td>Data used in the BH3 indicator come from a range of sources, comprising both open-source data and information which is commercially sensitive.

Please see below a summary of input data currently used in the indicator:
Fishing pressure (Open):
https://www.ices.dk/sites/pub/Publication%20Reports/Advice/2021/Special_Requests/ospar.2021.11.pdf

Commercial aggregate extraction pressure (Open & commercially sensitive):
Data forthcoming following publication of OSPAR QSR 2023. Please note, EMS & AIS may not be shared publicly due to commercial sensitivity.

OSPAR-scale habitat map (Open):
Data forthcoming following publication of OSPAR QSR 2023.

OSPAR Threatened & Declining Habitats (Open):
https://jncc.gov.uk/our-work/marine-habitat-data-product-ospar-threatened-andor-declining-habitats/

Species records (Open):
Data forthcoming following publication of OSPAR QSR 2023

Species sensitivity (Open):
Data forthcoming following publication of OSPAR QSR 2023.

Habitat sensitivity (Open):
https://www.marlin.ac.uk/

Sensitivity methods & aggregation to be published following publication of OSPAR QSR 2023.</td></tr>
<tr><td>Methodology</td><td>Parameters determined from biological data (e.g., Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)</td><td>Biotope Extent & Habitat Distribution
BH3 uses biological data to develop a map that shows the extent and distribution of habitats (based on survey and modelled data), including the mapped extent of any relevant features (e.g., records and distribution of species and biotopes, such as EUNIS Level 6 habitats or other biological characteristics). Biotopes are defined as a combination of an abiotic habitat and its associated community of species.

Habitat Sensitivity
BH3 assesses the sensitivity of key structural, functional, and characterizing species of benthic habitats in relation to a defined intensity of a given pressure. Sensitivity is assessed in terms of resistance (ability to withstand pressure) and resilience (ability to return to an unimpacted state, using key biological data, such as biological traits (e.g., life history and the ecology of the key and characterizing species). Please see below a summary of the types of species used to assess sensitivity.

<table>
<tr><th>Category</th><th>Description</th></tr>
<tr><td>Key structural species</td><td>The species provides a distinct habitat that supports an associated community. Loss/degradation of this species population would result in loss/degradation of the associated community.</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><td></td><td></td><td>
<table>
<tr><td>Key functional species</td><td>Species that maintain community structure and function through interactions with other members of that community (for example through predation, or grazing). Loss/degradation of this species population would result in rapid, cascading changes in the community.</td></tr>
<tr><td>Important characteristic species</td><td>Species characteristic of the biotope (dominant, and frequent) and important for the classification of the habitat. Loss/degradation of these species populations may result in changes in habitat classification.</td></tr>
</table>

Species Sensitivity
In addition to habitat sensitivity information, BH3 considered the sensitivity of in-situ sampled species records. Sensitivity is assessed using resistance and resilience and is assessed for key ecological groupings to identify species characteristic of sublittoral sediment habitats and sublittoral rock habitats. Bray Curtis similarity measure was used to quantify habitat characterizing species similarity, based on trait expression and habitat preference. Furthermore, ordination analysis (non-metric Multidimensional Scaling; nMDS) and cluster analyses (hierarchical agglomerative using group averaging) were used to identify ecological groups to assess sensitivity.
</td></tr>
<tr><td></td><td>Parameters determined from pressure data
(e.g., total SAR, years not fished, trawling interval)</td><td>
BH3 currently assesses potential physical disturbance from fishing and commercial aggregate extraction, although, the method can be adapted for wider activities known to cause physical pressure, where pressure and sensitivity data are available. Additional activities will be considered in forthcoming assessments.

Fishing pressure:
Total annual SAR values
Range SAR assessments, accounting for interannual SAR variability (please see BH3 CEMP Guidelines for further detail on fishing variability assessments). Range assessments are conducted across OSPAR QSR & MSFD intervals.
Areas not fished

Commercial aggregate extraction pressure:
Licensed extraction areas, with associated extraction statistics (where available).
Commercial aggregate dredging footprint and dredging intensity (gridded) as either:
The total volume dredged per licensed extraction area/per grid cell, or.
The total duration of extraction in units of time and volume dredged per grid cell as gridded spatial data, indicative of the activity intensity, including both vessel Automatic Identification System (AIS) and Electronic Monitoring System (EMS) data.
Total annual SAR values calculated on 50 x 50 m grid for dredge footprint & intensity.
Confirmation of OSPAR Contracting Parties where aggregate extraction activity is known not to occur.
Please note, the BH3 method can be adapted for wider activities known to cause physical pressure, where data are available and sensitivity information enables calculations of potential disturbance. Additional activities will be considered in forthcoming assessments.
</td></tr>
<tr><td></td><td>Algorithm type
List of categorical information (Presence/Absence, …)
Direct measurements (counts, areas, concentrations, …)
Single or multimetric indicators using basic arithmetic
Indicators using multivariate and complex statistics
Indicators derived from modelling approaches
Indicators reporting on trends</td><td>
2. Biological data, derived from survey/in-situ species records (presence only) when assessing species sensitivity; faunal data from surveys are used for biotope classification to inform habitat maps.

5. Modelling approaches (GIS/spatial models) developed in R, Python, ArcPy for ESRI ArcGIS and the ESRI GIS Model Builder.

Data science methods are employed in Python and R for assessments of sensitivity and the development of summary statistics & data visualizations.
</td></tr>
</table>

	References for state-pressure relation	Pressure-activity links The selection of pressures assessed in BH3 is informed by the JNCC Pressures Activities Database, an information system used by the statutory regulatory bodies in the UK for the management of anthropogenic pressure in marine environments. The selection of physical pressures considered relevant to assessed activities is informed by literature review, which provides an evidence base for understanding the relationships between human activities and their associated pressures in marine environments. The outputs of the literature review used to identify relevant pressures, where a pressure-activity relationship is known, can be found in the JNCC Marine Pressures Activities Database (version 1.5 used in QSR 2023 assessment) (Robson et al., 2018). Please see Defra (2015) and Robson et al., (2018) for further detail on pressure activity relationships with associated justifications, confidence assessments and risk profiles indicative of the likelihood of a given activity causing a particular pressure. Pressure-activity references: British Marine Aggregate Producers Association (BMAPA) & The Crown Estate (TCE), 2017. The impacts of marine aggregate dredging. Good Practice Guidance: Extraction by Dredging of Aggregates from England's Seabed. 9-10. [Online] Available at: https://www.bmapa.org/documents/BMAPA_TCE_Good_Practice_Guidance_04.2017.pdf (Accessed April 2021). Defra, (2015). Validating an Activity-Pressure Matrix. Report R.2435. [Online] Available at: http://randd.defra.gov.uk/Default.aspx?Menu=Menu&Module=More&Location=None&Compl eted=0&ProjectID=19471 (Accessed April 2022). Church N.J., Carter A.J., Tobin D., Edwards D., Eassom A., Cameron A., Johnson G.E., Robson, L.M. & Webb K.E., (2016). JNCC Recommended Pressure Mapping Methodology 1. Abrasion: Methods paper for creating a geo-data layer for the pressure 'Physical Damage (Reversible Change) - Penetration and/or disturbance of the substrate below the surface of the seabed, including abrasion'. JNCC report No. 515, JNCC, Peterborough. Cooper K.M., Eggleton J.D., Vize S.J., Vanstaen K., Smith R., Boyd S.E., Ware S., Morris C.D., Curtis, M. l., Limpenny D.S. & Meadows W.J., (2005). Assessment of the re-habilitation of the seabed following marine aggregate dredging-part II. Cefas Science Series Technical Report No. 130. Cefas Lowestoft. 82. Desprez, M., Stolk, A., and Cooper, K.M. (2022). Marine aggregate extraction and the Marine Strategy Framework Directive: A review of existing research. ICES Cooperative Research Reports, Vol. 354. 64 pp. https://doi.org/10.17895/ices.pub.19248542 Eigaard Ole R., Bastardie F., Breen M., Dinesen G. E. 1, Hintzen N. T., Laffargue P., Mortensen L. O., Nielsen J. R., Nilsson H. C., O'Neill F. G., Polet H., Reid D. G., Sala A., Skold M., Smith C., Sørensen T. K., Tully O., Zengin M., and Rijnsdorp A. D., (2015). Estimating seafloor pressure from demersal trawls, seines, and dredges based on gear design and dimensions. ICES Journal of Marine Science. Foden, J., Rogers, S. I. & Jones, A. P., (2010). Recovery of UK seafloor habitats from benthic fishing and aggregate extraction - towards a cumulative impact assessment. Marine Ecology Progress Series, 411, 259–270. Foden, J., Rogers, S. I. & Jones, A. P., (2011). Human pressures on UK seafloor habitats: a cumulative impact assessment. Marine Ecology Progress Series, 428, 33–47. ICES (2016) Effects of extraction of marine sediments on the marine environment 2005– 2011. ICES Cooperative Research Report No. 330. 206 pp ISBN 978-87-7482-179-3 ISSN 1017-6195 Available at: https://doi.org/10.17895/ices.pub.5498 ICES (2019a). Workshop to evaluate and test operational assessment of human activities causing physical disturbance and loss to seabed habitats (MSFD D6 C1, C2 and C4) (WKBEDPRES2). ICES Scientific Reports. 1:69. 87 pp. http://doi.org/10.17895/ices.pub.5611 ICES. (2019b). Working Group on the Effects of Extraction of Marine Sediments on the Marine Ecosystem (WGEXT). ICES Scientific Reports. 1:87. 133 pp. http://doi.org/10.17895/ices.pub.5733 Jennings, S., Alvsvag, J., Cotter, A. J., Ehrish, S., Greenstreet, S. P., Jarre-Teichmann, A., et al., (1999). Fishing effects on the northeat Atlantic shelf seas:patterns in fishing effort, diversity and community structure. III. International trawling effort in the North Sea: an analysis of spatial and temporal trends. Fisheries Research, 40, 125-134. Jennings, S., Lee, J., & Hiddink, J. G., (2012). Assessing fishery footprints and trade-offs between landings value, habitat sensitivity, and fishing impacts to inform marine spatial planning and an ecosystem approach. ICES Journal of Marine Science, 1-11. JNCC, (2011). Review of methods for mapping anthropogenic pressures in UK waters in support of the Marine Biodiversity Monitoring R&D Programme. Briefing paper to UKMMAS evidence groups. Presented 06/10/2011.

Korpinen, S., Meski, L., Andersen, J. H., & Laamanen, M., (2012). Human pressures and their potential impact on the baltic sea ecosystem. Ecological Indicators, 15, 105-114.
Last, K. S., Hendrick, V. J., Beveridge, C. M., Davies, A. J. (2011). Measuring the effects of suspended particulate matter and smothering on the behaviour, growth and survival of key species found in areas associated with aggregate dredging. pp. 70
Newell, R. C., Seiderer, L. J., Hitchcock, D. R. (1998). The impact of dredging works on coastal waters: A review of the sensitivity to disturbance and subsequent recovery of biological resources on the seabed. Oceanography and Marine Biology. 36, 127-78.
Newell, R. C., and Woodcock, T. A. (2013). Aggregate dredging and the marine environment: an overview of recent research and current industry practice. The Crown Estate. 165pp ISBN: 978-1-906410-41-4
OSPAR Commission, (2009). Summary assessment of sand and gravel extraction in the OSPAR maritime area. Publication number 434/2009. https://www.ospar.org/documents?v=7149
OSPAR Commission (2011). Pressure list and descriptions. Paper to ICG-COBAM 11/8/1 Add.1-E (amended version 25th March 2011) presented by ICG-Cumulative Effects. OSPAR Commission, London.
OSPAR Commission, (2014). OSPAR Joint Assessment and Monitoring Programme (JAMP) 2014-2023. [Online] Available at: http://www.ospar.org/documents?d=32988. (Accessed 10/11/2021)
OSPAR Commission, (2017a). OSPAR Intermediate Assessment 2017. OSPAR Commission. London.
OSPAR Commission, (2021). Feeder Report 2021 – Extraction of non-living Resources. Version 1.0.0. [Online]. Available at: https://oap.ospar.org/en/ospar-assessments/quality-status-reports/qsr-2023/other-assessments/extraction-non-living-resources/ (Accessed 28/04/2022).
Robson, L.M., Fincham, J., Peckett, F.J., Frost, N., Jackson, C., Carter, A.J. & Matear, L., (2018). UK Marine Pressures-Activities Database "PAD": Methods Report, JNCC Report No. 624, JNCC, Peterborough.
Robson, L.M., Fincham, J., Peckett, F.J., Frost, N., Jackson, C., Carter, A.J. & Matear, L. (2018). UK Marine Pressures-Activities Database "PAD": Methods Report. JNCC Report No. 624, JNCC, Peterborough, ISSN 0963-8091.
Schroeder, A., L. Gutow & M. Gusky (2008). FishPact. Auswirkungen von Grundschleppnetzfischereien sowie von Sand- und Kiesabbauvorhaben auf die Meeresbodenstruktur und das Benthos in den Schutzgebieten der deutschen AWZ der Nordsee (MAR 36032/15). Report for the Bundesamt für Naturschutz.
The Crown Estate, (2021). Introduction. Electronic Monitoring System Annual Report 2020. [Online] Available at: https://www.thecrownestate.co.uk/media/3995/2021-ems-report.pdf (Accessed April 2022)
Tillin, H. M., Houghton, J. Saunders, E., Drabble, R., and Hull, S. C. (2011) Direct and Indirect Impacts of Aggregate Dredging. Science Monograph Series No. 1. Marine ASLF. 41pp.
ISBN: 978 0 907545 43 9.

Sensitivity:
BH3 uses sensitivity information derived from the Marine Evidence-based Sensitivity Assessment (MarESA). Pressures considered in BH3 assessments are derived from definitions developed by the OSPAR Intersessional Correspondence Group on Cumulative Effects (ICG-C) and assessed against set sensitivity benchmarks, developed for indicating changes in a given receptor (e.g., habitat or species), in response to an assessed pressure.

Please see the following link for references & methodology on assessing MarESA sensitivity: https://www.marlin.ac.uk/sensitivity/sensitivity_rationale

JNCC (2011). Review of methods for mapping anthropogenic pressures in UK waters in support of the Marine Biodiversity Monitoring R&D Programme. Briefing paper to UKMMAS evidence groups. Presented 06/10/2011.
Last, E.K., Matear, L. & Robson, L.M., (2020). Developing a method for broadscale & feature-level sensitivity assessments: the MarESA aggregation. JNCC Report No. 662, JNCC, Peterborough, ISSN 0963-8091.
Tillin, H.M., Hull, S.C. & Tyler-Walters, H (2010). Development of a Sensitivity Matrix (pressures-MCZ/MPA features). Defra Contract No. MB0102 Task 3A, Report No. 22. http://jncc.defra.gov.uk/pdf/MB0102_Sensitivity_Assessment%5B1%5D.pdf
Tillin, H., Tyler-Walters, H., (2014a). Assessing the sensitivity of subtidal sedimentary habitats to pressures associated with marine activities. Phase 1 Report: Rationale and proposed ecological groupings for Level 5 biotopes against which sensitivity assessments would be best undertaken. JNCC Report No. 512A, 68 pp. Available from http://jncc.defra.gov.uk/page-6790

		Tillin, H. & Tyler-Walters, H., (2014b). Assessing the sensitivity of subtidal sedimentary habitats to pressures associated with marine activities. Phase 2 Report – Literature review and sensitivity assessments for ecological groups for circalittoral and offshore Level 5 biotopes. JNCC Report No. 512B, 260 pp. Available from: http://jncc.defra.gov.uk/PDF/Report 512-B_phase2_web.pdf Tillin, H.M. & Tyler-Walters, H., (2015). Finalised list of definitions of pressures and benchmarks for sensitivity assessment. 7-9. Tyler-Walters, H., Tillin, H.M., d'Avack, E.A.S., Perry, F. & Stamp, T., (2018). Marine Evidence based Sensitivity Assessment (MarESA) – A Guide. Plymouth: Marine Biological Association. Zacharias, M. A., and Gregr, E. J. (2005). Sensitivity and Vulnerability in Marine Environments: an Approach to Identifying Vulnerable Marine Areas. Conservation Biology. 19, 86-97.		
	Uncertainty estimation methodology	Outputs are developed with accompanying confidence maps to indicate uncertainty in component data layers used in assessments. Please see an example from the OSPAR Intermediate Assessment 2017 for reference (Results Extended - Confidence Section). https://oap.ospar.org/en/ospar-assessments/intermediate-assessment-2017/biodiversity-status/habitats/extent-physical-damage-predominant-and-special-habitats/ Sensitivity assessments are also completed with accompanying evaluations of the evidence used when defining receptor sensitivity against a given pressure. These confidence assessments consider the quality of the evidence used (e.g., peer-reviewed literature), the applicability of the evidence used (e.g., is it from a direct study on the assessed feature), and the degree of concordance (e.g., do underpinning evidence agree/align with the concluded sensitivity). Please see the following MarESA link for further detail on this approach - Table 6. Confidence assessment categories for evidence: https://www.marlin.ac.uk/sensitivity/sensitivity_rationale		
	Coding availability (e.g., scripts, GitHub)	Code underpinning the BH3 assessments is currently in development and will be made available following the publication of the OSPAR QSR 2023; scripts will be made available via GitHub, developed in the following languages: R Python ArcPy for ESRI ArcGIS SQL		
	Threshold present	Extent thresholds currently under discussion at OSPAR level; national thresholds are already used in DE and UK: DE: A broad habitat type is in GES, when at least 10% of its area is permanently without physical disturbance AND the area which is highly disturbed (disturbance categories 5-9) is less than 25% of the total habitat area. UK: The level of exposure to pressure at the level of the MSFD sub-regions should not result in more than Moderate Impact/vulnerability of the habitat (disturbance categories 0-4) (dependent on the sensitivity of the habitat to this pressure). Percentage values were calculated for each sub-region, to compare against a threshold of 15%, which was set as a potential indicator target in 2012.		
	Threshold methodology	Proposal by benthic experts, public consultation, and policy approval process		
Output	Output variable type	☐ Continuous	⌖ Categorical	⌖ Proportional
	Output variable range / classes	Outputs are expressed in spatial formats, with maps and GIS layers produced for sensitivity, pressure, and disturbance. In addition, summary statistics, such as the proportion of habitats under varying levels of disturbance are calculated and presented in tabular and graphical formats.		
	Output availability (e.g., report, website, reference)	BH3 assessment outputs are available online, from locations dependent on the application of the indicator: OSPAR Intermediate Assessment 2017: https://oap.ospar.org/en/ospar-assessments/intermediate-assessment-2017/biodiversity-status/habitats/extent-physical-damage-predominant-and-special-habitats/ OSPAR Quality Status Report 2023: Forthcoming UK Marine Strategy: https://moat.cefas.co.uk/biodiversity-food-webs-and-marine-protected-areas/benthic-habitats/		
	Uncertainty handling (e.g., present confidence interval)	Outputs are developed with accompanying confidence maps to indicate uncertainty in component data layers used in assessments. Please see an example from the OSPAR Intermediate Assessment 2017 for reference (Results Extended - Confidence Section). https://oap.ospar.org/en/ospar-assessments/intermediate-assessment-2017/biodiversity-status/habitats/extent-physical-damage-predominant-and-special-habitats/ Sensitivity assessments are also completed with accompanying evaluations of the evidence used when defining receptor sensitivity against a given pressure. These confidence assessments consider the quality of the evidence used (e.g., peer-reviewed literature), the applicability of the evidence used		

		(e.g., is it from a direct study on the assessed feature), and the degree of concordance (e.g., do underpinning evidence agree/align with the concluded sensitivity). Please see the following MarESA link for further detail on this approach - Table 6. Confidence assessment categories for evidence: https://www.marlin.ac.uk/sensitivity/sensitivity_rationale
	Spatial resolution (e.g., grid cell size, habitat level)	BH3 assessments are undertaken at the greatest possible level of detail, governed by the resolution of available habitat, sensitivity, and pressure information. For assessments of fishing pressure, the spatial resolution of assessments ranges from Broad Habitat-scale to biotope resolution (e.g., EUNIS Level 6), within a VMS square (0.05 decimal degree grid). For assessments of commercial aggregate extraction, the spatial resolution of assessments ranges from Broad Habitat-scale to biotope resolution (e.g., EUNIS Level 6), within a 50 m x 50 m projected grid. Please note, the resolution of outputs may vary, should new activities be included in future work.
	Temporal resolution	Assessment intervals are governed by the underlying legislative drivers of assessment. For example, OSPAR-scale assessments are undertaken at decadal intervals for the Quality Status Reports, whereas MSFD and UK Marine Strategy assessments are completed on 6-year intervals.
	Seabed habitat levels presented?	BH3 assessments are presented at the highest possible resolution of detail available in habitat maps, ranging from biotope-scale (e.g., EUNIS Level 6) to Broadscale Habitats (EUNIS) & Broad Habitat Types (MSFD).
More info	Indicator lead person	Cristina Vina-Herbon, Liam Matear, Axel Kreutle, Petra Schmitt
	Indicator data contact	Cristina Vina-Herbon, Liam Matear, Axel Kreutle, Petra Schmitt
	References / Literature / Project websites	Please see the OSPAR website for a full overview of evidence base used with BH3. Data used in the QSR 2023 are forthcoming.

OSPAR: BH4

Indicator description	Indicator name	Area of habitat loss (OSPAR BH4)		
	Indicator description	The indicator estimates the extent and proportion of benthic habitats that is lost due to human activities. Habitat loss can be caused by sealed loss (e.g. placement of structures, disposal of sediments), by unsealed loss (permanent change of the sediment type due to e.g. bottom trawling or aggregate extraction) or biogenic loss (historic loss of biogenic substrate). Currently a pilot assessment is developed that focuses on some activities causing sealed and unsealed loss in the North Sea. Sealed loss is assessed by spatially combining the footprint of the structure with the extent and distribution of benthic habitat types. The outcome is the area lost per habitat type and sub-region in % and km². For unsealed loss a risk assessment is produced, where the probability of habitat change is combined with the intensity of the pressure. The probability of habitat change is determined by the substrate type and the energy at the seabed, taking also recoverability into account. The outcome is the area and proportion of habitat in different risk categories.		
	Type of indicator	X Model	☐ Empirical-based	☐ Pressure
	Pressure assessed	Physical loss		
	Human activity	In general, all activities causing habitat loss with adequate data. The pilot assessment focuses on oil and gas platforms, pipelines, offshore wind farms, bottom trawling and aggregate extraction.		
	MSFD criteria /descriptor	D6C1, D6C4		
	How does the indicator relate to benthic biological diversity?	The indicator does not directly relate to benthic community parameters, but habitat loss affects benthic diversity and structure and function.		
	How does the indicator relate to benthic community structure and function?	The indicator does not directly relate to benthic community parameters, but habitat loss affects benthic diversity and structure and function.		
	Indicator status	X Under development	☐ Applied for MSFD	☐ Applied for other management, if so, for what:
	Regions with operational assessments	Pilot assessment is produced for the MFSD sub-region Greater North Sea, including the Kattegat and the English Channel / OSPAR Region II (North Sea). The pilot assessment is part of the coming OSPAR QSR.		

Input data	Biological data input (e.g. monitoring program, time series, sampling method)	-			
	Targeted organisms	☐ Infauna	☐ Epifauna	☐ Demersal fish	☐ Other: ….
	Environmental data input (e.g. empirical/modelled, source, time series)	Data derived from EUSeaMap comprise a suite of EMODnet environmental datasets, including EMODnet Bathymetry, EMODnet Geology and Copernicus marine services via the Copernicus Marine Environment Monitoring Service (CMEMS). Additional physical data used for the calculation of the models include data on light attenuation, light at the seabed and kinetic, current and wave energy datasets.			
	Pressure data input (e.g. time series, empirical/modelled, source, national/international)	For pilot assessment: Distribution and extent of offshore structures VMS data (SAR values) Data on aggregate extraction Other activities causing loss can be included as well, if spatial data are available. For activities causing unsealed loss, information on the intensity of the activity is needed.			
	Data availability	OWF, oil and gas platforms: ODIMS, EMODnet Pipelines: EMODnet VMS: ICES (2021) Aggregate extraction: OSPAR data call, EMODnet Data availability for other activities (e.g. dumping sites, capital dredging, coastal defense structures) and historic distribution of biogenic reefs not yet sufficient			
Methodology	Parameters determined from biological data (e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)	-			
	Parameters determined from pressure data (e.g. total SAR, years not fished, trawling interval)	Spatial footprint of sealed loss by offshore structures VMS: mean SAR value per assessment period (6 years) Aggregate extraction: impact of dredging related to method and intensity of extraction (not yet determined due to data limitations) Please note, the BH4 method can be adapted for wider activities known to cause physical loss, where data are available and sensitivity information enable calculations of potential habitat loss. Additional activities will be considered in forthcoming assessments.			
	Algorithm type List of categorical information (Presence/Absence, …) Direct measurements (counts, areas, concentrations, …) Single or multimetric indicators using basic arithmetics Indicators using multivariate and complex statistics Indicators derived from modelling approaches Indicators reporting on trends	5			
	References for state-pressure relation	Bottom trawling: Oberle F.K.J., Swarzenski P.W., Reddy C.M., Nelson R.K., Baasch B. & Hanebuth T.J.J. (2016): Deciphering the lithological consequences of bottom trawling to sedimentary habitats on the shelf. Journal of Marine Systems 159:120–131 Mengual B., Cayocca F., Le Hir P., Draye P., Laffargue P., Vincent B. & Garlan T. (2016): Influence of bottom trawling on sediment resuspension in the 'Grande-Vasière' area (Bay of Biscay, France). - Ocean Dynamics 1181-1207 Schratzberger M. & Jennings S. (2002): Impacts of chronic trawling disturbance on meiofaunal communities. - Marine Biology 141 (5): 991-1000 Aggregate extraction:			

		Foden J., Rogers S.I. & Jones A.P. (2009): Recovery rates of UK seabed habitats after cessation of aggregate extraction, Mar. Ecol. Prog. Ser., 390, 15–26 Desprez M. (2012): Synthèse bibliographique. L'impact des extractions de granulats marins sur les écosystèmes marins et la biodiversité. Les études de l'UNPG - Nature et paysage. Mielck F., Michaelis R., Hass H.C., Hertel S., Ganal C. & Armonies W. (2021): Persistent effects of sand extraction on habitats and associated benthic communities in the German Bight. Biogeosciences, 18, 3565–3577 Newell R.C. & Woodcock T.A. (2013): Aggregate Dredging and the Marine Environment: an overview of recent research and current industry practice. The Crown Estate. Tillin H.M., Houghton A.J., Saunders J.E., Drabble R. & Hull S.C. (2011): Direct and Indirect Impacts of Aggregate Dredging. Marine ALSF Science Monograph Series No.1. MEPF 10.		
	Uncertainty estimation methodology	Confidence assessment for data and methodology		
	Coding availability (e.g. scripts, GitHub)	-		
	Threshold present	-		
	Threshold methodology	-		
Output	Output variable type	x Continuous	x Categorical	x Proportional
	Output variable range / classes	The output for sealed loss is given as km² and proportion of area lost per habitat type. The output for unsealed loss is the habitat area and proportion per risk category.		
	Output availability (e.g. report, website, reference)	OSPAR Quality Status Report 2023 (forthcoming)		
	Uncertainty handling (e.g. present confidence interval)	Descriptive confidence assessment		
	Spatial resolution (e.g. grid cell size, habitat level)	VMS grid size: 0.05 x 0.05° Habitat level: EUNIS (2019) level 3		
	Temporal resolution	Sealed loss: all structures currently present were assessed Trawling: 6 year-assessment periods (2009-2014, 2015-2020) Aggregate extraction: reporting year 2019		
	Seabed habitat levels presented?	EUNIS 2019 classification is used, this is compatible to MSFD broad habitat types		
More info	Indicator lead person	Axel Kreutle, Petra Schmitt, Cristina Vina-Herbon, Liam Matear		
	Indicator data contact			
	References / Literature / Project websites	-		

OSPAR: BISI

Indicator description	Indicator name	Benthic Indicator Species Index (BISI)		
	Indicator description	The Benthic Indicator Species Index (BISI) is a methodology to compare combined occurrence (densities, presence-absence or biomass) of a set of area/habitat specific indicator species (with specific indicator value; i.e. sensitive for specific disturbances characteristic for specific habitat or representing important ecosystem functions and sufficient common under good quality conditions to monitor with realistic efforts) with estimated reference occurrence indicating the relative quality status of the benthic habitat(s) (reflected in General BISI value), with indication of the most probable pressure(s) leading to impact as reflected by the quality status and expected effects on ecosystem functioning (as indicated by Specific BISI values based on sub-selections of indicator species with specific sensitivities, characteristics, functions).		
	Type of indicator	☐ Model	X Empirical-based	☐ Pressure
	Pressure assessed	Basically, the general quality status (result of all pressures that are at stake) is assessed (reflected in General BISI value). However specific assessments give evidence on the relative importance of specific disturbances leading to the observed quality status. Applied in areas with physical disturbance (seafloor disturbing fisheries, eutrophication and increased levels of pollutants ('ecological disturbance'), changes in hydro-morpho-dynamics due to suppletion, and combinations of pressures).		
	Human activity	Demersal fisheries/bottom trawling or organic enrichment and/or inorganic pollutants or		

		sediment suppletion or extraction or disturbances due to artificial structures (e.g. wind turbines, platforms) leading to hydro-morphological changes.			
	MSFD criteria /descriptor	D6 (D6C5) (Has been used for D6C3 in the Netherlands as well)			
	How does the indicator relate to benthic biological diversity?	The BISI basically is a type of benthic diversity indicator; as especially sensitive and characteristic species and species with important ecological functions are selected, the assessment results will relate to total (taxonomic) diversity and presence and diversity of different traits and functions as well.			
	How does the indicator relate to benthic community structure and function?	The BISI basically is a type of benthic diversity indicator; as especially sensitive and characteristic species and species with important ecological functions are selected, the assessment results will relate to total (taxonomic) diversity and presence and diversity of different traits and functions as well.			
	Indicator status	X Under development	X Applied for MSFD	X Applied for other management, if so, for what: - Benthic habitat quality assessment for the Habitat Directive (HD) - Evaluation of effectivity of measures (i.e. closed areas for seafloor disturbing fisheries compared to open areas) - Effects of and short-term recovery after pilot sand nourishment - Test-runs for broad-scale transnational application for OSPAR	
	Regions with operational assessments	Areas of Dutch EEZ (MSFD, HD, effectivity of measures ‘closed areas’, suppletion case study) Application in effect study pilot nourishment (Amelander Zeegat) BISI assessment sheets (sets of indicator species with reference occurrences) developed for HD-areas Dutch transitional waters (Wadden Sea – H1110, H1140, H1130); Western Scheldt - H1130); Eastern Scheldt - H1160) Testing application around international Dogger Bank and application for (entire) Greater North Sea region within frame of OSPAR in progress.			
Input data	Biological data input (e.g. monitoring program, time series, sampling method)	Benthic community data of any kind. - Application in Dutch North Sea (MSFD, HD, effectivity of measures) based on combined boxcore and benthic dredge data (each having their own indicator species), for Cleaver Bank based on Hamon grab and video transects; also, separate assessments based on singular sample technique data. - Application pilot nourishment, and test application and application in progress within frame of OSPAR (matching MSFD requirements) based on grabs / boxcores of about similar size (~0.1 m2). Evaluation of separate years (might allow trend analysis) or assessment periods dependent on data availability. It is suggested that identified references (e.g. for BHTs at Assessment Unit / Marine Reporting Unit scale related to monitoring technique), are re-used for assessments of the same kind. References for composite areas (consisting of several BHTs in case of combined evaluation can be composed surface-area based, or taking the number of samples per BHT into account). For derivation of references for new regions/ habitats/ techniques/ etc, comprehensive datasets and data from low pressure areas in particular, are needed. Although different types of observation data can potentially be used, so far BISI assessment sheets have been developed predominantly making use of densities for indicator species, and occasionally relative presence (portion of samples where indicator species present) is used (like in case of colony-forming species).			
	Targeted organisms	X Infauna	X Epifauna	☐ Demersal fish	☐ Other: ….

	Environmental data input (e.g. empirical/modelled, source, time series)	Variable – typically habitat maps are used (e.g. BHTs), but also areas distinguished on certain environmental conditions determining type of communities or species distributions could be used.
	Pressure data input (e.g. time series, empirical/modelled, source, national/international)	In principle, no pressure data are used; basically, BISI provides an overall (general) quality status assessment of the benthic habitats, with an indication of what might be the dominant pressure(s) at stake, leading to impact resulting in the observed status. Pressure data can be of use for interpretation and validation. In case no reference values for the BISI of concern (alternative regions/ habitats/ techniques/ etc have been derived before, information on pressure distribution is needed to identify low pressure areas (those low pressure areas should be low pressure area with regards to all potential (most important) pressures that might be at stake (e.g. representative SAR data, nutrient/ pollutant levels, etc).
	Data availability	- Application in Dutch North Sea: Benthic community data Dutch EEZ (MSFD, HD, effectivity of measures): https://www.informatiehuismarien.nl/uk/open-data-viewer/ select 'KRM bevroren monitoring data IHM' – 'D6C5 conditie van het benthisch habitat'. Assessment areas: https://www.informatiehuismarien.nl/open-data-viewer/?opendatafolder=Beleid en beheer&opendatalayer=KRM gebieden Habitat layer: EUSeaMap (2019). Broad-Scale Predictive Habitat Map – EUNIS classification. EMODnet broad-scale seabed habitat map for Europe (v2019), licensed under CC-BY 4.0 from the European Marine Observation and Data Network (EMODnet) Seabed Habitats initiative (www.emodnet-seabedhabitats.eu), funded by the European Commission. - Application pilot nourishment: Benthic community data pilot nourishment: https://www.informatiehuismarien.nl/open-data-viewer/?opendatafolder=Ecologie&opendatalayer=Macrobenthos (KG2)
Methodology	Parameters determined from biological data (e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)	BISI-score; which is actually a type of indicator species diversity, based on combined observed occurrence divided by reference occurrence for selected species. In specific assessments to identify most important pressure(s) having impact and resulting in the observed quality status, and most important effects of observed quality status on ecological functioning, indicator values (weights) are given to indicator species to distinguish in relative indicator value. BISI-scores are calculated at the level of Assessment Units x Habitats (e.g. BHTs) providing average BISI values ± pooled standard deviation. Reference occurrences are extracted from comprehensive low pressure datasets taking maximum observed occurrences into account (expert judgement on the representativity of data like are they really from low pressure areas, obtained with similar techniques, in similar habitats, sufficient data for several years etc, is involved in defining the reference occurrence).
	Parameters determined from pressure data (e.g. total SAR, years not fished, trawling interval)	In principle no pressure data are used, except for pressure data to identify low pressure areas for reference setting. In that case pressure maps (other project outputs) are used; e.g. average SAR or subSAR per c-square and modelled nutrient/pollutant distributions.
	Algorithm type List of categorical information (Presence/Absence, …) Direct measurements (counts, areas, concentrations, …) Single or multimetric indicators using basic arithmetics Indicators using multivariate and complex statistics Indicators derived from modelling approaches Indicators reporting on trends	3 Although calculations combining several indicator species (potentially using weights; indicator values) with a general assessment and several specific assessments might look complex, the basic formula is relative simple: BISI = exp((1/S)*∑ln(IVi*(Oi/Ri))) BISI = Benthic Indicator Species Index; S = Number of indicator species included; IVi = Species specific Indicator Value calculated as species specific standard indicator value ivi (value between 0-1) divided by average indicator value ivavg; Oi = Observed occurrence species i (either presence/absence ratio, density or biomass); Ri = Reference occurrence species i (presence/absence ratio, density or biomass under reference conditions). ('exp' is similar to putting e to the power of the formula as indicated, which equals the inverse natural logarithm, as a back-transformation of the natural logarithm (ln) taken from the occurrence-to-reference ratios). It has to be noticed that above formula refers to BISI v2, that first applications (applications in Dutch North Sea) were according to v1: BISI = exp((1/S)∑(IVi)log(Oi/Ri)), so with not

		matching log-transformation and back-transformation and indicator values brought outside the log (nevertheless although this gives deviating BISI-results, t is still the case that good quality habitats score high and poor quality habitats low. Moreover, at present v3 is in development where the indicator values are again placed outside the log-term: $\text{BISI} = \exp((1/S)*\sum(IVi*\ln(Oi/Ri)))$, where suggested indicator values are halved when below '1' (both to improve the distinction among specific assessments), and where the methodology to calculated pooled standard deviation has been improved/ corrected. Therefore, it can be concluded that the current methodology (although in use) is still in developmental phase and therefore not published in a scientific journal yet. This is planned after finalization of test cases (International Dogger Bank), applications with additional data (new MSFD-cycle) for the areas of the Dutch EEZ, and applications for the Greater North Sea region at BHT level; indicative at the end of 2022.
	References for state-pressure relation	Evidence for state-pressure relations is extracted from scientific literature and species trait databases at the individual indicator species level (used to select indicator species per assessment unit, habitat, technique and specific assessment, and to define indicator value (relative weight is the calculation of the BISI). References per species provided in: Wijnhoven, S. (2019d). Assessment tool: 'Benthic Indicator Species Index (BISI)': Application of BISI v2 in soft sediment habitats of OSPAR region II (Greater North Sea region). v311219. (see 'PotIndSpec' sheet). Wijnhoven, S. (2019b). Assessment tool 'Benthic Indicator Species Index (BISI)': Application of BISI v2 for marine Habitat Directive habitat types of the Dutch 'Delta-waters', the Wadden Sea and the coastal zone of the North Sea. v070120. (see sheets starting with 'Ref ecotopes ….').
	Uncertainty estimation methodology	Calculation of BISI scores comes with pooled standard deviations, used in testing of potential differences in BISI-scores (representative for quality status benthic habitats) between cases or in time. At present responsiveness of BISI to differences in numbers of samples, sampled surface area, type of technique, number of indicator species included in BISI, relative difference in indicator values (weight factor used in calculation), calculating BISI for entire area or combining results for subdivisions, is being investigated in the international Dogger Bank test case. Results (discussed in OSPAR Benthic Habitat Expert Group, but not published yet) are used to fine-tune (increase sensitivity and comparability) and update methodology to BISI v3 (as indicated above).
	Coding availability (e.g. scripts, GitHub)	Methodology elaborated for a variety of systems and habitats with specific indicator species and reference values related to type of samples. Entire indices (BISIs) available in Assessment tools (in Excel) where average occurrences (densities, presence-ratio), standard deviations and number of samples can be filled in and result in a general and specific BISI scores (see manual included in Assessment tools): Wijnhoven, S. (2019d). Assessment tool: 'Benthic Indicator Species Index (BISI)': Application of BISI v2 in soft sediment habitats of OSPAR region II (Greater North Sea region). v311219. Wijnhoven, S. (2019c). Assessment tool: 'Benthic Indicator Species Index (BISI)': Application of BISI v2 in the Dutch North Sea with consolidation of earlier identified references. v311219. Wijnhoven, S. (2019b). Assessment tool 'Benthic Indicator Species Index (BISI)': Application of BISI v2 for marine Habitat Directive habitat types of the Dutch 'Delta-waters', the Wadden Sea and the coastal zone of the North Sea. v070120. Wijnhoven, S. (2017b). Assessment tool 'Benthic Indicator Species Index (BISI)'; Application of BISI v1 in the Dutch North Sea areas of evaluation. v260917. Appendix 2 of Wijnhoven & Bos (2017). (Assessment tools not available in coding yet).
	Threshold present	No
	Threshold methodology	-
Output	Output variable type	X Continuous ☐ Categorical ☐ Proportional
	Output variable range / classes	0.01 (very poor habitat quality status; no indicator species present, or occurrences indicator species always 100 times lower than the reference occurrence) to 1 (good habitat quality status, i.e. at reference level; all indicator species present in about reference occurrences;

		in case indicator species are missing, occurrences of others should be above reference) (In theory BISI-scores can increase to 100 (range from 0.1-100 and has a logarithmic scale); in those cases habitat quality is considered to be at reference level (BISI=1), but exceedance is uncommon).
	Output availability (e.g. report, website, reference)	Wijnhoven, S. (2018a). T0 beoordeling kwaliteitstoestand NCP op basis van de Benthische Indicator Soorten Index (BISI). Toestand en ontwikkelingen van benthische habitats en KRM gebieden op de Noordzee in en voorafgaand aan 2015. Rapport Ecoauthor & Wageningen Marine Research. Ecoauthor Report Series 2018 – 01, Heinkenszand, the Netherlands (in Dutch). KRM Factsheet D6C3 (2018). KRM factsheet D6C3 Benthische habitats kwaliteit (BISI). Onderdeel van de Mariene Strategie (deel 1) (Min IenW & Min LNV, 2018). Tevens Annex 1 van Wijnhoven (2018a) (in Dutch). Janssen, J.A.M. (ed.), R.J. Bijlsma (ed.), G.H.P. Arts, M.J. Baptist, S.M. Hennekens, B. de Knegt, T. van der Meij, J.H.J. Schaminée, A.J. van Strien, S. Wijnhoven, T.J.W. Ysebaert (2020). Habitats Directive Report 2019: Annex D Habitat Types – Background Document. Statutory Research Tasks Unit for Nature & the Environment, WOt Technical Report 171. 97 p. (https://edepot.wur.nl/514490; in Dutch). Wijnhoven, S. (2021). Korte-termijn-effecten pilotsuppletie Amelander Zeegat. Analyse ontwikkeling benthische habitats met behulp van de BISI. Ecoauthor Report Series 2021 - 01, Heinkenszand, the Netherlands. (In Dutch).
	Uncertainty handling (e.g. present confidence interval)	Calculation of BISI scores comes with pooled standard deviations, used in testing of potential differences in BISI-scores (representative for quality status benthic habitats) between cases or in time.
	Spatial resolution (e.g. grid cell size, habitat level)	Assessments are done at the level of assessment units or Broad Habitats within assessment units (these can be Marine Reporting Units/ Subregions, (National) part of North Sea, (parts of) MPAs, specific basins or marine waters, etc, where BHTs or ecotopes are assessed individually or in combined assessments for composite areas).
	Temporal resolution	Typically, a result is obtained for a certain moment in time (campaign, month, season, year) but can also an assessment period (e.g. 6-year MSFD period or OSPAR review period). A series of observations provides information on benthic habitat quality status (as indicated by BISI-scores) development, e.g. using trend analyses and/or for specific cases in a BACI approach.
	Seabed habitat levels presented?	Typically, the Broad Habitat Type is the level of assessment, but in other cases combinations of BHTs are used for composite areas. In transitional waters, ecotopes are used as the level of assessment. In a nourishment testcase, alternative (more detailed habitats/ecotopes were used). (In principle it can be any level of habitat, when sufficient data are (or have been) available (from other cases) to identify reference levels (sufficient data from low pressure areas are needed).
More info	Indicator lead person	Sander Wijnhoven; sander.wijnhoven@ecoauthor.net
	Indicator data contact	Sander Wijnhoven; sander.wijnhoven@ecoauthor.net
	References / Literature / Project websites	http://ecoauthor.net/bisi/

ICES: L1

Indicator description	Indicator name	L1		
	Indicator description	The indicator assumes that a population is affected by trawling if animals are disturbed by trawls during their life span. Only species in the community with a longevity less than the average interval between two successive trawling events will not be affected.		
	Type of indicator	X Model	☐ Empirical-based	☐ Pressure
	Pressure assessed	Physical abrasion		
	Human activity	Demersal fisheries / bottom trawling		
	MSFD criteria /descriptor	D6 and D1		
	How does the indicator relate to benthic biological diversity?	The indicator assumes that integrity of the seabed habitat is compromised when species cannot complete their full life cycle without being disturbed by trawl gear. This is a very precautionary assumption, which will benefit rare and sensitive species and, as such, diversity.		
	How does the indicator relate to benthic community structure and function?	Unclear.		

	Indicator status	X Under development	☐ Applied for MSFD	☐ Applied for other management, if so, for what:	
	Regions with operational assessments	Baltic Sea and North Sea (+ all ICES areas where a longevity prediction is available).			
Input data	Biological data input (e.g. monitoring program, time series, sampling method)	Variable – Benthic box core data, or benthic epifauna data obtained with trawls. Benthic species longevity information			
Input data	Targeted organisms	X Infauna	X Epifauna	☐ Demersal fish	☐ Other: ….
Input data	Environmental data input (e.g. empirical/modelled, source, time series)	Variable - typically continuous environmental conditions (e.g. % mud and gravel, tidal bed shear stress) but can be habitat classes			
Input data	Pressure data input (e.g. time series, empirical/modelled, source, national/international)	International Swept area ratio per year, VMS-based.			
Input data	Data availability	North Sea - Underlying benthic and environmental data is available here: https://doi.org/10.5061/dryad.th2c5f7 Baltic Sea - Underlying benthic and environmental data is available here https://github.com/Dvandenderen/Baltic-benthic-status			
Methodology	Parameters determined from biological data (e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)	Distribution of biomass over longevity classes.			
Methodology	Parameters determined from pressure data (e.g. total SAR, years not fished, trawling interval)	Annual average (surface) Swept Area Ratio per grid cell			
Methodology	Algorithm type List of categorical information (Presence/Absence, …) Direct measurements (counts, areas, concentrations, …) Single or multimetric indicators using basic arithmetics Indicators using multivariate and complex statistics Indicators derived from modelling approaches Indicators reporting on trends	5			
Methodology	References for state-pressure relation	Unavailable			
Methodology	Uncertainty estimation methodology	Indicator needs a biomass-longevity distribution, which is derived from a statistical model. Uncertainty in the biomass-longevity distribution can be obtained from the statistical model and used to estimate uncertainty in the impact score, as done in Rijnsdorp et al. 2020 (https://doi.org/10.1093/icesjms/fsaa050)			
Methodology	Coding availability (e.g. scripts, GitHub)	https://github.com/ices-eg/WKTRADE3			
Methodology	Threshold present	No			
Methodology	Threshold methodology	None			
Output	Output variable type	X Continuous	☐ Categorical	☐ Proportional	
Output	Output variable range / classes				
Output	Output availability (e.g. report, website, reference)	https://github.com/ices-eg/WKTRADE3			
Output	Uncertainty handling (e.g. present confidence interval)	Output did not handle uncertainty in ICES WKTRADE3 work. Confidence intervals are presented in Rijnsdorp et al. (2020) (https://doi.org/10.1093/icesjms/fsaa050)			
Output	Spatial resolution (e.g. grid cell size, habitat level)	Grid cell size (0.05 x 0.05 as well as 1 x 1 minute)			

	Temporal resolution	Annual, output can use averages over multiple years
	Seabed habitat levels presented?	MSFD broad habitat types
More info	Indicator lead person	Adriaan Rijnsdorp: adriaan.rijnsdorp@wur.nl
	Indicator data contact	Daniël van Denderen: pdvd@aqua.dtu.dk
	References / Literature / Project websites	Rijnsdorp et al., 2020; https://doi.org/10.1093/icesjms/fsaa050

ICES: L2

Indicator description	Indicator name	L2			
	Indicator description	The indicator estimates the decrease in median longevity in response to trawling. Median longevity is the longevity where 50% of the community biomass is above/below. The decrease is based on a statistical relationship between trawling intensity and benthic longevity from the North Sea.			
	Type of indicator	☐ Model	X Empirical-based	☐ Pressure	
	Pressure assessed	Physical abrasion			
	Human activity	Demersal fisheries / bottom trawling			
	MSFD criteria /descriptor	D6 and D1			
	How does the indicator relate to benthic biological diversity?	Low values of the indicator compared to reference conditions imply a reduction in the proportion of long-lived species in the community and therefore potential loss of diversity.			
	How does the indicator relate to benthic community structure and function?	L2 method incorporates information on structure by estimating the biomass-longevity composition. The indicator is less likely to be a good indicator of function.			
	Indicator status	X Under development	☐ Applied for MSFD	☐ Applied for other management, if so, for what:	
	Regions with operational assessments	North Sea			
Input data	Biological data input (e.g. monitoring program, time series, sampling method)	Benthic box core and/or grab data Benthic species longevity information			
	Targeted organisms	X Infauna	X Epifauna	☐ Demersal fish	☐ Other: ….
	Environmental data input (e.g. empirical/modelled, source, time series)	Empirical (and modeled) information on sediment type. Modeled information on tidal wave stress.			
	Pressure data input (e.g. time series, empirical/modelled, source, national/international)	International / subsurface swept area ratio per year, VMS-based.			
	Data availability	Underlying benthic and environmental data is available here: https://doi.org/10.5061/dryad.th2c5f7			
Methodology	Parameters determined from biological data (e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)	Distribution of biomass over longevity classes.			
	Parameters determined from pressure data (e.g. total SAR, years not fished, trawling interval)	Annual average (subsurface) Swept Area Ratio per grid cell			
	Algorithm type List of categorical information (Presence/Absence, …) Direct measurements (counts, areas, concentrations, …) Single or multimetric indicators using basic arithmetics Indicators using multivariate	6 (spatial correlations, not time trends)			

	and complex statistics Indicators derived from modelling approaches Indicators reporting on trends			
	References for state-pressure relation	Rijnsdorp et al. 2018 https://doi.org/10.1002/eap.1731		
	Uncertainty estimation methodology	Pressure-state relationship is derived from a statistical model. Uncertainty can be obtained from the statistical model and used to estimate uncertainty in the impact score, as done in Rijnsdorp et al. 2020 (https://doi.org/10.1093/icesjms/fsaa050)		
	Coding availability (e.g. scripts, GitHub)	No		
	Threshold present	No		
	Threshold methodology	None		
Output	Output variable type	X Continuous	☐ Categorical	☐ Proportional
	Output variable range / classes			
	Output availability (e.g. report, website, reference)	No, but can be reproduced using the Rijnsdorp et al. 2018 https://doi.org/10.1002/eap.1731 paper and ICES WGFBIT output		
	Uncertainty handling (e.g. present confidence interval)	Yes - 95% confidence limits are obtained from the statistical model		
	Spatial resolution (e.g. grid cell size, habitat level)	Grid cell size (0.05 x 0.05 as well as 1 x 1 minute)		
	Temporal resolution	Annual, output can use averages over multiple years		
	Seabed habitat levels presented?	Yes, old EUNIS-3 habitat types (similar to the MSFD broad habitat types that are currently used)		
More info	Indicator lead person	Adriaan Rijnsdorp: adriaan.rijnsdorp@wur.nl		
	Indicator data contact	Daniël van Denderen: pdvd@aqua.dtu.dk		
	References / Literature / Project websites	Rijnsdorp et al. 2018 https://doi.org/10.1002/eap.1731 Rijnsdorp et al. 2020 https://doi.org/10.1093/icesjms/fsaa050		

ICES: PD

Indicator description	Indicator name	PD		
	Indicator description	The PD method is a mechanistic model that is based on the logistic population growth equation, which is generally applied in ecology and fisheries to describe how populations change in size in response to exploitation. The model needs depletion (d) and recovery (r) parameters, which were estimated from all globally available trawl impact studies for infauna and epifauna. The method and its parameter estimates are therefore applicable globally. In the PD method, the recovery rate of a community depends on the longevity distribution of an untrawled community. In the WKBENTH and WGFBIT report, the PD method was applied to the North Sea, and the longevity distribution of an untrawled community was estimated based on a statistical model The response variable presented by the PD method is the relative benthic biomass (RBS), which is the whole community benthic biomass relative to carrying capacity (i.e. the sum of the biomass of fauna of all different longevities relative to what it would have been with no fishing).		
	Type of indicator	X Model	☐ Empirical-based	☐ Pressure
	Pressure assessed	Bottom trawling fisheries		
	Human activity	Fisheries		
	MSFD criteria /descriptor	D6C3 D6C5		
	How does the indicator relate to benthic biological diversity?	RBS as estimated by the PD method incorporates information on the total biomass, which relates closely to functioning of ecosystems, and the relative abundance of different longevity classes, which relates to the structure and biodiversity.		
	How does the indicator relate to benthic community structure and function?	RBS as estimated by the PD method incorporates information on the total biomass, which relates closely to functioning of ecosystems, and the relative abundance of different longevity classes, which relates to the structure and biodiversity.		
	Indicator status	X Under development	☐ Applied for MSFD	☐ Applied for other management, if so, for

					what:
	Regions with operational assessments	North Sea and Baltic are complete, with development for the rest of all European Seas ongoing in WGFBIT and likely to be complete with 1-2 years			
Input data	Biological data input (e.g. monitoring program, time series, sampling method)	Spatial datasets of benthic communities by genus with biomass over environmental gradients, including stations with no or minimal trawling. Longevity trait categorization by genus (default option is <1y, 1-3y, 3-10y, >10y).			
	Targeted organisms	X Infauna	X Epifauna	☐ Demersal fish	☐ Other: ….
	Environmental data input (e.g. empirical/modelled, source, time series)	Environmental data layers are needed to match the biological samples and fit statistical models of the biomass-longevity distribution. These models are combined with the data layers to create the sensitivity layer.			
	Pressure data input (e.g. time series, empirical/modelled, source, national/international)	Bottom trawling swept-area-ratio by metier, derived from VMS and logbooks			
	Data availability	Fully available for North Sea and Baltic Sea. Environmental layers available for all areas. Bottom trawling swept-area-ratio availability varies between areas, but some effort layer available for most areas.			
Methodology	Parameters determined from biological data (e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)	Biomass by genus.			
	Parameters determined from pressure data (e.g. total SAR, years not fished, trawling interval)	Bottom trawling swept-area-ratio by metier.			
	Algorithm type List of categorical information (Presence/Absence, …) Direct measurements (counts, areas, concentrations, …) Single or multimetric indicators using basic arithmetics Indicators using multivariate and complex statistics Indicators derived from modelling approaches Indicators reporting on trends	Indicators derived from modelling approaches			
	References for state-pressure relation	Hiddink, J.G., Jennings, S., Sciberras, M., Szostek, C.L., Hughes, K.M., Ellis, N., Rijnsdorp, A.D., McConnaughey, R.A., Mazor, T., Hilborn, R., Collie, J.S., Pitcher, R., Amoroso, R.O., Parma, A.M., Suuronen, P. & Kaiser, M.J. (2017) Global analysis of depletion and recovery of seabed biota following bottom trawling disturbance. Proceedings of the National Academy of Sciences, 114, 8301–8306. Hiddink, J.G., Kaiser, M.J., Sciberras, M., McConnaughey, R.A., Mazor, T., Hilborn, R., Collie, J.S., Pitcher, R., Parma, A.M., Suuronen, P., Rijnsdorp, A.D. & Jennings, S. (2020) Selection of indicators for assessing and managing the impacts of bottom trawling on seabed habitats. Journal of Applied Ecology, 57, 1199-1209. Sciberras, M., Hiddink, J.G., Jennings, S., Szostek, C.L., Hughes, K.M., Kneafsey, B., Clarke, L.J., Ellis, N., Rijnsdorp, A.D., McConnaughey, R.A., Hilborn, R., Collie, J.S., Pitcher, C.R., Amoroso, R.O., Parma, A.M., Suuronen, P. & Kaiser, M.J. (2018) Response of benthic fauna to experimental bottom fishing: a global meta-analysis. Fish and Fisheries, 19, 698-715.			
	Uncertainty estimation methodology	Has been developed in FBIT based on bootstrapping method.			
	Coding availability (e.g. scripts, GitHub)	https://github.com/ices-eg/FBIT			
	Threshold present	Been explored in a manuscript 'Setting thresholds for good marine ecosystem state and significant adverse impacts' by Jan Geert Hiddink, Sebastian Valanko, Adam J. Delargy, Daniel van Denderen			

<table>
<tr><td></td><td>Threshold methodology</td><td colspan="3">Staying inside the range of natural variation</td></tr>
<tr><td rowspan="8">Output</td><td>Output variable type</td><td>☐ Continuous</td><td>☐ Categorical</td><td>X Proportional</td></tr>
<tr><td>Output variable range / classes</td><td colspan="3">0 (community lost/destroyed) to 1 (unimpacted)</td></tr>
<tr><td>Output availability (e.g. report, website, reference)</td><td colspan="3">Available: https://github.com/ices-eg/FBIT</td></tr>
<tr><td>Uncertainty handling (e.g. present confidence interval)</td><td colspan="3">Confidence Interval presented of mean RBS value</td></tr>
<tr><td>Spatial resolution (e.g. grid cell size, habitat level)</td><td colspan="3">0.05 x 0.05 latitude x longitude degree</td></tr>
<tr><td>Temporal resolution</td><td colspan="3">Once; determining annual average over the used time period</td></tr>
<tr><td>Seabed habitat levels presented</td><td colspan="3">MSFD broad habitat types</td></tr>
<tr><td colspan="4"></td></tr>
<tr><td rowspan="3">More info</td><td>Indicator lead person</td><td colspan="3">Jan Geert Hiddink: j.hiddink@bangor.ac.uk</td></tr>
<tr><td>Indicator data person</td><td colspan="3">Daniël van Denderen: pdvd@aqua.dtu.dk</td></tr>
<tr><td>References / Literature / Project website</td><td colspan="3">ICES WGFBIT reports
Rijnsdorp et al., 2020; DOI:10.1093/icesjms/fsaa050</td></tr>
</table>

HELCOM: CumI

<table>
<tr><td rowspan="10">Indicator description</td><td>Indicator name</td><td colspan="3">Cumulative impact from physical pressures on benthic biotopes (CumI)</td></tr>
<tr><td>Indicator description</td><td colspan="3">Assessment of the potential cumulative impact from all relevant physical pressures on benthic biotopes, considering the pressures' extent, frequency and intensity and the sensitivity of the biotopes towards these pressures.

The assessment is done by taking a biotope map containing BHT polygons (broad habitat types), assigning a general sensitivity to it (towards physical pressures on the basis of the benthic communities that typical occur on the BHT) and a specific sensitivity towards bottom trawling. This information is overlayed with spatial information on physical pressures (extent, intensity, frequency) to arrive at one impact map per pressure. In a last step these impacts maps are overlayed to derive the cumulative impact when these pressures act on the same place at the same time. The result is the cumulative impact.</td></tr>
<tr><td>Type of indicator</td><td>X Model</td><td>☐ Empirical-based</td><td>☐ Pressure</td></tr>
<tr><td>Pressure assessed</td><td colspan="3">Abrasion/penetration/extraction/disposal of sediment, sealing and smothering, general physical disturbance and removal/loss of sediment/substrate/habitat</td></tr>
<tr><td>Human activity</td><td colspan="3">bottom trawling fishery and mariculture, extraction and disposal of sediments (e. g. dredging and dumping), construction/building and operation of pipelines and cables, platforms and wind farms, coastal protection and shipping</td></tr>
<tr><td>MSFD criteria /descriptor</td><td colspan="3">D6C3 (the indicator also results in additional information that can be used in D6C4, but that MSFD criterion is not the specific target of the indicator)</td></tr>
<tr><td>How does the indicator relate to benthic biological diversity?</td><td colspan="3">The indicator does not directly target biological diversity. However, indirectly, diversity is included. Every physical pressure on the seafloor directly affects the biological diversity when it reaches a level that makes species or biotopes disappear. Alpha diversity is already affected earlier without entire species/biotopes disappearing from an area in terms of beta/gamma diversity.</td></tr>
<tr><td>How does the indicator relate to benthic community structure and function?</td><td colspan="3">Physical pressures directly affect the abundance of species, the abiotic structure of the biotope and thus lead to changes in community structure. The nature and strength of these changes are directly depending on the sensitivity of the underlying biotopes towards the physical pressures. A changed structure subsequently leads to altered functions if the disturbance is too strong.</td></tr>
<tr><td>Indicator status</td><td>☐ Under development</td><td>x Applied for MSFD</td><td>x Applied for other management, if so, for what: HELCOM holistic assessment</td></tr>
<tr><td>Regions with operational assessments</td><td colspan="3">Baltic Sea</td></tr>
<tr><td>Input data</td><td>Biological data input (e.g. monitoring program, time series, sampling method)</td><td colspan="3">None, apart from a (static) biotope map derived from abiotic and biological data (including species distribution data and species/community sensitivity)</td></tr>
</table>

	Targeted organisms	x Infauna	x Epifauna	☐ Demersal fish	☐ Other: ….
	Environmental data input (e.g. empirical/modelled, source, time series)	Empirical or modelled biotope map			
	Pressure data input (e.g. time series, empirical/modelled, source, national/international)	Time series, modelled and source data on pressure distribution, intensity (e.g. SAR values, shipping density, footprints of cables/pipelines, outlines of sediment extraction areas and deposition sites)			
	Data availability	Data are mostly from HELCOM Data and Maps Service, layers preprocessed for CumI are available at https://portal.helcom.fi/workspaces/EN-BENTHIC-191/default.aspx?RootFolder=%2Fworkspaces%2FEN%2DBENTHIC%2D191%2FShared%20Documents%2FCumI%2FCumI%2DR%2FData			
Methodology	Parameters determined from biological data (e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)	Biotope sensitivity (resistance and resilience)			
Methodology	Parameters determined from pressure data (e.g. total SAR, years not fished, trawling interval)	Areas of potential impact (spatially buffered point and line data) divided into four or less intensity zones, derived from the raw pressure data			
Methodology	Algorithm type List of categorical information (Presence/Absence, …) Direct measurements (counts, areas, concentrations, …) Single or multimetric indicators using basic arithmetics Indicators using multivariate and complex statistics Indicators derived from modelling approaches Indicators	3 & 5			

	reporting on trends			
	References for state-pressure relation	Crain et al. 2008, DOI: 10.1111/j.1461-0248.2008.01253.x Hiddink et al. 2017, DOI: 10.1073/pnas.1618858114 HELCOM 2010, DOI: 10.13140/RG.2.1.2148.6961 ICES 2019, Workshop to valuate and test operational assessments of human activities causing physical disturbance and loss to seabed habitats (MSFD D6 C1, C2 and C4) (WKBEDPRES2), 30 September – 2 October 2019, Copenhagen, Denmark. ICES Scientific Reports. 1:69. 87 pp HELCOM (2016) Synthesis of the impacts of human activities on seabed habitats WP 3_1 BalticBOOST WS-1-2016 Copenhagen: 7pp Korpinen et al. 2013, DOI: 10.1016/j.marpolbul.2013.06.036 Ware et al. 2009, 10.1016/j.marpolbul.2009.08.031		
	Uncertainty estimation methodology	Currently an expert judgement, based on quality of underlying pressure data Numerical uncertainty assessment in planning phase		
	Coding availability (e.g. scripts, GitHub)	The CumI is available at https://github.com/torstenberg/CumI (currently version 1.1 from 2021-09-08)		
	Threshold present	Yes, quality threshold present		
	Threshold methodology	Biological valuation: threshold is set where adverse effects begin (sensu MSFD)		
Output	Output variable type	☐ Continuous	x Categorical	☐ Proportional
	Output variable range / classes	One of 6 disturbance classes: very low, low, moderate 1, moderate 2, moderate 3, high In addition one category for cumulative disturbance leading to loss (not part of CumI assessment result, but transferred as input data to D6C4)		
	Output availability (e.g. report, website, reference)	Current version of indicator report available at https://portal.helcom.fi/workspaces/EN-BENTHIC-191/Shared%20Documents/CumI/CumI%20indicator%20report/Cumulative-impact-indicator-report-2022-04-13.docx		
	Uncertainty handling (e.g. present confidence interval)	Currently, global 'medium' confidence applied		
	Spatial resolution (e.g. grid cell size, habitat level)	The indicator uses exact polygons when available. Raster data from e.g. fishery have a spatial resolution of 0.05 x 0.05 degrees (geographical coordinates) corresponding to roughly 3 x 5 km		
	Temporal resolution	Current assessment on basis of data from 2011–2016 (will be update to 2016–2021 with an updated indicator in 2022)		
	Seabed habitat levels presented?	MSFD broad habitat types (BHT)		
More info	Indicator lead person	Torsten Berg: berg@marilim.de		
	Indicator data contact	Torsten Berg: berg@marilim.de		
	References / Literature / Project websites	indicator report: https://portal.helcom.fi/workspaces/EN-BENTHIC-191/Shared%20Documents/CumI/CumI%20indicator%20report/Cumulative-impact-indicator-report-2022-04-13.docx		

HELCOM: BQI

Indicator description	Indicator name	HELCOM State of the Soft-bottom macrofauna community			
	Indicator description	BQI evaluation against regionally agreed threshold values. Spatial coverage is not complete in the Baltic Sea			
	Type of indicator	☐ Model	X Empirical-based	☐ Pressure	
	Pressure assessed	All relevant cumulative pressures on these habitat types.			
	Human activity	All relevant cumulative activities occurring on these habitat types.			
	MSFD criteria /descriptor	Expected to be a contributor to D6C5. Also applied under D5C8 to support eutrophication assessment.			
	How does the indicator relate to benthic biological diversity?	Direct monitoring and sampling of soft-bottom macrofauna communities to evaluate state (station based assessment extrapolated to HELCOM assessment units and awarded a confidence based on the spatial/temporal coverage of sampling.			
	How does the indicator relate to benthic community structure and function?	Direct evaluation of soft-bottom macrofauna community and structure at monitoring stations.			
	Indicator status	☐ Under development	x Applied for MSFD	x Applied for other management, if so, for what: WFD in coastal areas for some Contracting Parties	
	Regions with operational assessments	Large areas of the Baltic Sea are covered by the indicator, see latest report from 2018. Spatial coverage expected to remain the same in 2022/2023 assessment as ongoing work has not been completed in southerly/south-easterly assessment units to date.			
Input data	Biological data input (e.g. monitoring program, time series, sampling method)	National monitoring via regionally agreed monitoring programme. Grab sampling with spatially distributed station samples and time series data sets. See page 17 of indicator report.			
	Targeted organisms	x Infauna	☐ Epifauna	☐ Demersal fish	☐ Other: ….
	Environmental data input (e.g. empirical/modelled, source, time series)	Empirical data from national monitoring.			
	Pressure data input (e.g. time series, empirical/modelled, source, national/international)	NA			
	Data availability	See data section in indicator report, page 18.			
Methodology	Parameters determined from biological data (e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)	Benthic Quality Index (BQI), where the abundance weighted proportion of sensitive to tolerant taxa and the diversity of the community are the determining parameters. See page 10 of report.			
	Parameters determined from pressure data (e.g. total SAR, years not fished, trawling interval)	NA			
	Algorithm type List of categorical information (Presence/Absence, …) Direct measurements (counts, areas, concentrations, …) Single or multimetric indicators using basic arithmetics Indicators using multivariate and complex	BQI calculation presented on page 10 of report.			

	statistics Indicators derived from modelling approaches Indicators reporting on trends			
	References for state-pressure relation	The BQI approach has been developed through several consecutive studies (Rosenberg et al. 2004, Leonardsson et al. 2009, Leonardsson et al. 2015, Leonardsson et al. 2016 and Blomqvist & Leonardsson 2016).		
	Uncertainty estimation methodology	Calculate the 20th percentile of the stored 100 000 mean BQI-values. In order to account for spatial, temporal and sample replicate imbalance a bootstrap procedure was used to estimate the 20th percentile to be compared against the threshold value. The 20th percentile is used as a precautionary or "fail-safe" approach (Carstensen 2007, Leonardsson et al. 2009) placing results of high uncertainty into lower status categories. The evaluation is awarded a confidence based on the input data and statistical outputs, see page 5 of report.		
	Coding availability (e.g. scripts, GitHub)	https://github.com/helcomsecretariat/StateOfTheSoftbottomMacrofaunaCommunity		
	Threshold present	Sub-basin specific threshold values are applied in the indicator (the sub-basins being ecologically relevant management divisions of the Baltic Sea). These are described and presented in a table on page 8 and 9 of the indicator report.		
	Threshold methodology	In Bothnian Bay, The Quark, Bothnian Sea, Åland Sea, Northern Baltic Proper and Western Gotland Basin, where the method follows Leonardsson et al. (2009), the Swedish intercalibrated BQI good-moderate threshold values, developed for outer coastal waters under the EU Water Framework Directive, are considered to also be applicable for the open sea assessment units. The establishment of these threshold values is based on both statistical tests and expert judgment, using data from areas without local disturbance to define high and good status as baselines, as described in Leonardsson et al. (2009). In Gulf of Finland, Gulf of Riga, Eastern Gotland Basin, Bay of Mecklenburg and Kiel Bay the species sensitivity values used for calculation of BQI follows Schiele et al. (2016). In this method the described fauna sub-sets that occur in the assessment unit are first identified. Threshold values are then calculated for each subset according to a pragmatic statistical scheme developed by Perus et al. (2007) and later modified during an intercalibration process, as described by Carletti & Heiskanen (2009). In short, this method sets threshold values as 0.6 times the 10th percentile of the top 10 % of all index values within a subset.		
Output	Output variable type	☐ Continuous	x Categorical	☐ Proportional
	Output variable range / classes			
	Output availability (e.g. report, website, reference)	Latest version of indicator report: https://helcom.fi/wp-content/uploads/2019/08/State-of-the-soft-bottom-macrofauna-community-HELCOM-core-indicator-2018.pdf		
	Uncertainty handling (e.g. present confidence interval)	The evaluation is awarded a confidence based on the input data and statistical outputs, see page 5 of report.		
	Spatial resolution (e.g. grid cell size, habitat level)	Stations results are used to carry out evaluations at regionally agreed and ecologically relevant HELCOM assessment units.		
	Temporal resolution	Stations are sampled annually via national monitoring and the indicator is applied for 6-year periods to evaluate in a harmonious way with MSFD requirements.		
	Seabed habitat levels presented?	The indicator is only applicable to a merge of soft habitat types (mud, sand, and certain mixed habitat types). In sub-basins where a permanent halocline exist, the indicator is only applied above the halocline (i.e. areas <60 m deep).		
More info	Indicator lead person	Via HELCOM Secretariat – Owen Rowe (owen.rowe@helcom.fi) Expert leads under HELCOM - Henrik Nygård (Finland) and Mats Blomqvist (Sweden).		
	Indicator data contact	Via HELCOM Secretariat – Joni Kaitaranta (joni.kaitaranta@helcom.fi)		
	References / Literature / Project websites	Latest version of indicator report (where the literature references can be found): https://helcom.fi/wp-content/uploads/2019/08/State-of-the-soft-bottom-macrofauna-community-HELCOM-core-indicator-2018.pdf		

HELCOM: Condition of benthic habitats

Indicator description	Indicator name	Condition of benthic habitats			
	Indicator description	The indicator utilizes area, extent and quality (status) of habitat types to create an evaluation of overall benthic habitat status.			
	Type of indicator	☐ Model	X Empirical-based	☐ Pressure	
	Pressure assessed	Effects of multiple pressures or overall effects/condition.			
	Human activity	Impacts of multiple human activities on overall condition.			
	MSFD criteria /descriptor	D6C5			
	How does the indicator relate to benthic biological diversity?	Assessment procedure is composed of conditional classification of three different habitat properties: 1) area of the habitat, 2) extent of the habitat (range), and 3) quality of the habitat. For most of the elements assessment is based on comparison of current situation with reference level.			
	How does the indicator relate to benthic community structure and function?	The indicator utilizes either direct benthic monitoring data, indicators that represent relevant proxies for benthic status (e.g. water clarity), and in cases modeled data to evaluate areas, extent and quality that can be combined to provide an overall status or condition of benthic habitats.			
	Indicator status	x Under development	☐ Applied for MSFD	☐ Applied for other management, if so, for what:	
	Regions with operational assessments	Test cases in Estonian waters from circa 2017.			
Input data	Biological data input (e.g. monitoring program, time series, sampling method)	National monitoring data used in test cases.			
	Targeted organisms	☐ Infauna	☐ Epifauna	☐ Demersal fish	☐ Other: ….
	Environmental data input (e.g. empirical/modelled, source, time series)	Mixed data can be utilized.			
	Pressure data input (e.g. time series, empirical/modelled, source, national/international)	Relevant data would be from national monitoring.			
	Data availability	Some data that is suitable for the system is available from HELCOM and through national monitoring programmes.			
Methodology	Parameters determined from biological data (e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)	Evaluation of extent and quality.			
	Parameters determined from				

	pressure data (e.g. total SAR, years not fished, trawling interval)	
	Algorithm type List of categorical information (Presence/Absence, …) Direct measurements (counts, areas, concentrations, …) Single or multimetric indicators using basic arithmetics Indicators using multivariate and complex statistics Indicators derived from modelling approaches Indicators reporting on trends	Categorical with area based (5) evaluation.
	References for state-pressure relation	
	Uncertainty estimation methodology	
	Coding availability (e.g. scripts, GitHub)	
	Threshold present	
	Threshold methodology	
Output	Output variable type	☐ Continuous ☐ Categorical ☐ Proportional
	Output variable range / classes	
	Output availability (e.g. report, website, reference)	Test cases and 2017 version available online. https://portal.helcom.fi/meetings/STATE%20-%20CONSERVATION%207-2017-470/MeetingDocuments/3J-16%20Updated%20indicator%20report%20on%20condition%20of%20benthic%20habitats.pdf
	Uncertainty handling (e.g. present confidence interval)	
	Spatial resolution (e.g. grid cell size, habitat level)	Applied best at smallest grid level possible, dependent on data.
	Temporal resolution	Focused on 6 year assessment period, but some data series may be long term.
	Seabed habitat levels presented?	Biotope level or BHT level.
M	Indicator lead	Via HELCOM Secretariat – Owen Rowe (owen.rowe@helcom.fi)

	person	Expert leads under HELCOM – Gerog Martin
	Indicator data contact	No regional data is compiled for this currently, though data may be available for some aspects.
	References / Literature / Project websites	https://portal.helcom.fi/meetings/STATE%20-%20CONSERVATION%207-2017-470/MeetingDocuments/3J-16%20Updated%20indicator%20report%20on%20condition%20of%20benthic%20habitats.pdf This indicator as not been further developed since 2017. Processes in HELCOM have been awaiting TG Seabed guidance related to D6C5. Since the progress made under TG Seabed some further discussions have taken place in HELCOM that may be relevant to the pre-core indicator in this template (or interim approaches for considering D6C5) – for example: https://portal.helcom.fi/meetings/HOLAS3%20BENTHIC%201-2022-992/MeetingDocuments/2-1%20Summary%20of%20the%20latest%20developments%20of%20the%20HELCOM%20work%20on%20benthic%20habitats.pdf and https://portal.helcom.fi/meetings/EN-BENTHIC%207-2021-937/MeetingDocuments/4-2%20Aggregation%20and%20integration%20of%20spatial%20assessments.pdf. This pre-core indicator is not currently utilized regionally.

Member States: BEQI

Indicator description	Indicator name	Benthic Ecosystem Quality Index (BEQI)			
	Indicator description	This is the third level of the BEQI analyses, which evaluates the benthic macrofauna community per habitat. Threshold values defined for each parameter delimit condition classes wherein a characteristic benthic community is expected to occur. The BEQI evaluates the benthic community at the level of a habitat, rather than the evaluation of a single sample. It evaluated how much a habitat is changed compared to the reference habitat state (assessment – reference) or how much a human activity is changing the benthic habitat (impact – control).			
	Type of indicator	☐ Model	☐ Empirical-based	x Pressure	
	Pressure assessed	Evaluate the impact of any pressure based on dedicated impact-control monitoring data.			
	Human activity	Any human activity			
	MSFD criteria /descriptor	Descriptor 1 and 6.			
	How does the indicator relate to benthic biological diversity?	It includes a diversity parameter (number of species) for direct measurement of biological diversity.			
	How does the indicator relate to benthic community structure and function?	It includes a parameter reflecting community structure (species composition) and the total biomass, which relates closely to functioning of ecosystems.			
	Indicator status	☐ Under development	x Applied for MSFD	X Applied for other management, if so, for what: WFD and EIA processes	
	Regions with operational assessments	Southern North Sea (Belgian Waters)			
Input data	Biological data input (e.g. monitoring program, time series, sampling method)	Any biological data set consisting of impact-control or assessment-reference monitoring data.			
	Targeted organisms	x Infauna	x Epifauna	x Demersal fish	☐ Other: ….
	Environmental data input (e.g. empirical/modelled, source, time series)	No, but dataset needs to be collected under the same environmental condition (e.g. habitat) or dataset and evaluation need to be split in groups reflecting the same environmental conditions.			
	Pressure data input (e.g. time series, empirical/modelled, source, national/international)	Any pressure data (continuous, categorical) associated with the benthic dataset.			
	Data availability	The BEQI analyses, requires a certain amount of reference and assessment samples and sampling area per habitat. In this way the natural variability (spatial and temporal) is included.			
Methodology	Parameters determined from biological data (e.g. Species richness,	Species richness Bray Curtis similarity (species composition) Density			

	abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)	Biomass		
	Parameters determined from pressure data (e.g. total SAR, years not fished, trawling interval)	BEQI scores can be plotted against any pressure data variable.		
	Algorithm type List of categorical information (Presence/Absence, …) Direct measurements (counts, areas, concentrations, …) Single or multimetric indicators using basic arithmetics Indicators using multivariate and complex statistics Indicators derived from modelling approaches Indicators reporting on trends	4. indicator using complex statistics		
	References for state-pressure relation	A specific reference/control dataset is defined based on the type of evaluation (human activity impact or ecological state).		
	Uncertainty estimation methodology	Per evaluated parameter (species, similarity, density, biomass), a categorical estimation of uncertainty is given (		
	Coding availability (e.g. scripts, GitHub)	Yes; http://www.beqi.eu/		
	Threshold present	Yes, EQR of 0.6		
	Threshold methodology	The reference value of the good/moderate boundary (EQR 0.6) is determined based on the 5th percentile (number of species, similarity) or on the 2.5th and 97.5th percentile (density, biomass) out of the permutation distribution of each parameter of the control/reference dataset. The moderate/poor (EQR 0.4) and poor/bad (EQR 0.2) reference value were determined by equal scaling (respectively 2/3 or 4/3 and 1/3 or 5/3 of the good/moderate reference value), whereas the median value (number of species, similarity) or the 25th and 75th percentile (density, biomass) out of the permutation distribution was used as the reference value of the high/good boundary (EQR 0.8). The expected reference values for the BEQI parameters (as described above) are calculated per habitat from permutations executed over increased sampling surfaces. An algorithm was used that computed rarefaction curves using a random resampling procedure with replacement (i.e. bootstrapping, using 2000 random samples). This allows estimating, for any given sampling surface, the reference value that can be expected, which then can be compared with a similar sampling surface used to evaluate the current ecological status.		
Output	Output variable type	☐ Continuous	x Categorical	☐ Proportional
	Output variable range / classes	EQR: 0-1		
	Output availability (e.g. report, website, reference)	/ EQR score per parameter and an average EQR score per analysis		
	Uncertainty handling (e.g. present confidence interval)	Informative categorical score		
	Spatial resolution (e.g. grid cell size, habitat level)	Depending on data availability, but score per habitat type within an assessment area		
	Temporal resolution	Depending on data availability, but normally once (determining annual average over the used time period)		
	Seabed habitat levels presented?	Any seabed habitats, as long as assessment occurs per habitat type following a control-impact (or assessment – reference) data approach.		
More info	Indicator lead person	Gert Van Hoey		
	Indicator data contact	Gert.vanhoey@ilvo.vlaanderen.be		
	References / Literature / Project websites	Website with info and analyse software: http://www.beqi.eu/ Belgium MSFD assessment example:		

		https://odnature.naturalsciences.be/msfd/nl/assessments/2018/page-d1-d6 Van Hoey, G.; Drent, J.; Ysebaert, T.; Herman, P. (2007). The Benthic Ecosystem Quality Index (BEQI), intercalibration and assessment of Dutch coastal and transitional waters for the Water Framework Directive: Final report. NIOO Rapporten, 2007-02. NIOO. 244 pp. Van Hoey G., Seghers S., Festjens F., Dewitte B., Vanavermaete D., Jacobs L., Wittoeck J., Hillewaert H., Lefranc C., Vanhalst K., Hostens K., 2022. Influence of the disposal of dredged material on the marine sea-bottom ecosystem in the Belgian Part of the North Sea. ILVO-mededeling D/2022/03

Member States: DKI

Indicator description	Indicator name	DKI, the Danish quality index (DK: Dansk Kvalitetsindex)
	Indicator description	The DKI is a multi-metric state indicator developed for assessment of ecological quality of marine softbottom fauna in Danish territorial waters, including in the Western Baltic Sea, Belt seas and Øresund, Kattegat-Skagerrak, and the North Sea south to the Danish part of the Wadden Sea. The DKI (ver1) was developed for application using macrofaunal data retrieved from softbottom sediment samples collected by either a Van Veen grab (seabed area sampled: 0.1 m2) and or HAPS corer (seabed area sampled: 0.0143 m2). All samples are sieved at a mesh size of 1 mm. Faunal identification is carried out to the lowest taxonomical level possible, preferable to the species level. The Danish macrofaunal data informs an equation which comprises density by number of individuals (N), species density (S), diversity express as the Shannon-Wiener index (H' with log base 2), the H' value of an undisturbed reference condition (H'max) and AMBI categories (between 1-7 as a proxy for sensitivity-tolerance to nutrient loading and oxygen depletion) (Borja et al. 2007). The equation is: DKIver1 = (((1-(AMBI/7))+(H'/H'max))/2 * ((1-(1/N))+(1-(1/S)))/2) The DKIver1 was subsequently standardized for salinity (DKIver2) based on 2600 seabed samples sized ~0.1 m2 (as 1 Van Veen or 7 pooled HAPS corer) collected from 540 sites in April 2008. Maximum values of S and H' decreased and minimum values of AMBI increased with decreasing salinity from 28 to 8 psu at open water sites (e.g., Kattegat) but not in closed fjords or lagoons (Josefson 2008, Josefson 2009, Henriksen et al. (2014). When adjusting H'max and AMBImin relative to the salinity relation and omitting the 1-1/S factor, the modified DKI (ver2) equation is: DKIver2 = ((1 - ((AMBI-AMBImin/7)) + (H'/H'max))/2 * (1-(1/N) where: H'max = 2.117 + 0.086 * salinity, and AMBImin = 3.083 – 0.111 * salinity (Josefson 2008, Carstensen et al. 2014, Henriksen et al. 2014). In the past >10 years the DKIv2 has, however, been calculated based on only 5 pooled HAPS corer per site (due to further changes of monitoring programmes) resulting in lower DKI values (Henriksen et al. 2014). DKIver2 is used as a benthic macrofaunal indicator for GES assessment under the EU WFD (e.g., Hansen & Høgslund 2021). It has been suggested as a benthic indicator candidate for D6D3 (and D1) assessment. Gislason et al. (2017) found, however, DKIver2 to be significantly related to salinity but not to fishing pressure from bottom trawling measured as swept area ratio (SAR). The significant salinity response may be a result of the DKIver2 salinity standardization, which moreover, did not consider potential effects of differences in trawling intensity, nutrient loadings or frequency of hypoxia events that could also have affected density and species density at the reference sampling sites (Gislason et al. 2017). Based on linear mixed effect model, Gislason et al. (2017) found the log N (density of individuals) explained the largest proportion of variation (78%) as compared with BQI (Benthic Quality Index), DKI, AMBI, H' and log S (species density).

<table>
<tr><td rowspan="9"></td><td></td><td colspan="12">Furthermore, Gislason et al. (2017) found species density, S, substantially influence the variation in the DKIver2 indicator through the Shannon-Wiener index, H'. Species diversity and sensitivity indices depend on both balanced sampling designs and comparable estimates of species density across stations and years. Gislason et al. (2017) found high correlation between log N and log S. If density varies naturally between sites (or areas), or between years due to inter-annual variation in larval recruitment, the indices are likely to provide a variable background which prevents estimation of how species diversity respond to anthropogenic pressures (Gislason et al. 2017).</td></tr>
<tr><td>Type of indicator</td><td colspan="4">☐ Model</td><td colspan="4">X Empirical-based</td><td colspan="4">☐ Pressure</td></tr>
<tr><td>Pressure assessed</td><td colspan="12">Indirectly nutrient loading and derived oxygen depletion by including the AMBI eutrophication sensitivity-tolerance scores (categories: 1-7).</td></tr>
<tr><td>Human activity</td><td colspan="12">Agriculture (minor extent also industrial/household sewage)</td></tr>
<tr><td>MSFD criteria /descriptor</td><td colspan="12">D5C5 & C8,
D6C3 (& D1C3 for benthic habitats)</td></tr>
<tr><td>How does the indicator relate to benthic biological diversity?</td><td colspan="12">Comprise density (N), and species density (S) individually, and relation of the Shannon Wiener diversity index.</td></tr>
<tr><td>How does the indicator relate to benthic community structure and function?</td><td colspan="12">It does not.</td></tr>
<tr><td>Indicator status</td><td colspan="4">☐ Under development</td><td colspan="4">☐ Applied for MSFD</td><td colspan="4">X Applied for other management, if so, for what: WFD</td></tr>
<tr><td>Regions with operational assessments</td><td colspan="12">The Danish EEZ: Western Baltic Sea, Belt seas and Øresund, Kattegat-Skagerrak, North Sea south to the Danish Wadden Sea.</td></tr>
<tr><td rowspan="5">Input data</td><td>Biological data input (e.g. monitoring program, time series, sampling method)</td><td colspan="12">Time series from the Danish NOVANA monitoring programme (e.g., 2004-present) of benthic fauna (sites with replicate, grids with single samples) sampled using different sediment collecting gears.</td></tr>
<tr><td>Targeted organisms</td><td colspan="3">☐ Infauna</td><td colspan="3">☐ Epifauna</td><td colspan="3">☐ Demersal fish</td><td colspan="3">X Other: softbottom macrofauna (epifauna & infauna)</td></tr>
<tr><td>Environmental data input (e.g. empirical/modelled, source, time series)</td><td colspan="12">Salinity standardization (for DKIver2) based on data from 2008.</td></tr>
<tr><td>Pressure data input (e.g. time series, empirical/modelled, source, national/international)</td><td colspan="12">No.</td></tr>
<tr><td>Data availability</td><td colspan="12">See below.</td></tr>
<tr><td rowspan="3">Methodology</td><td>Parameters determined from biological data
(e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)</td><td colspan="12">Density of individuals, N (number of individuals)
Density of Species, S
Shannon-Wiener, H' and H'max
AMBI eutrofication sensitivity-tolerance classes 1-7, AMBI and AMBImin.</td></tr>
<tr><td>Parameters determined from pressure data
(e.g. total SAR, years not fished, trawling interval)</td><td colspan="12">Not considered.</td></tr>
<tr><td>Algorithm type
List of categorical information (Presence/Absence, …)
Direct measurements (counts, areas, concentrations, …)
Single or multimetric indicators using basic arithmetics
Indicators using multivariate and complex statistics
Indicators derived from modelling approaches</td><td colspan="12">2 and 3.</td></tr>
</table>

	Indicators reporting on trends	
	References for state-pressure relation	Not available.
	Uncertainty estimation methodology	Not available.
	Coding availability (e.g. scripts, GitHub)	AMBI software available at https://ambi.azti.es/ (Borja et al. 2000).
	Threshold present	The GES threshold (set as a border between god and moderate EcoQS G/M), was determined for DKIver2 by Josefson et al. (2009) as the value where faunal structure deterioration commenced, and identified from non-linear regression of DKIver2 and distance (pressure gradient from a urban sewage point source) in a Danish bay (Aarhus Bight). Based on this the 5th percentile value from GESmax was estimated using a bootstrap procedure, resulting in a G/M threshold for DKIver2 at 0.68 (Josefson et al. 2009).
	Threshold methodology	Proxy for pressure and statistical threshold from a single near-coastal area in inner Danish waters.
Output	Output variable type	☐ Continuous ☐ Categorical X Proportional
	Output variable range / classes	0-1, where 0 = no fauna and 1 = highest state.
	Output availability (e.g. report, website, reference)	Reports.
	Uncertainty handling (e.g. present confidence interval)	No.
	Spatial resolution (e.g. grid cell size, habitat level)	Point data (from grab or corer sediment samples) of 0.1 m2 seabed.
	Temporal resolution	Annual (April-May).
	Seabed habitat levels presented?	No.
More info	Indicator lead person	Jørgen L. S. Hansen, Danish Center for Environment and Energy (DCE), University of Aarhus, Denmark.
	Indicator data contact	See above.
	References / Literature / Project websites	Borja, A., Franco, J. & Perez, V. 2000: A marine biotic index to establish the ecological quality of soft bottom benthos within European estuarine and coastal environments. - Marine Pollution Bulletin 40, 1100-1114. Borja et al. 2007. An approach to the intercalibration of benthic ecological status assessment in the North Atlantic ecoregion, according to the European Water Framework Directive. Marine Pollution Bulletin 55, 42–52. Carstensen, J., Krause-Jensen, D. & Josefson, A. 2014. Development and testing of tools for intercalibration of phytoplankton, macrovegetation and benthic fauna in Danish coastal areas. Aarhus University, DCE – Danish Centre for Environment and Energy, 85 pp. - Scientific Report from DCE – Danish Centre for Environment and Energy No. 93. Hansen, J.W., Høgslund, S. (red.) 2021. Marine områder 2019. NOVANA. Aarhus Universitet, DCE - Nationalt Center for Miljø og Energi. Videnskabelig rapport fra DCE nr. 418. Henriksen et al. 2014. Danish contribution to the EU Water Framework Directive intercalibration phase 2. DCE - Danish Centre for Environment and Energy. Technical report from DCE nr. 37. Gislason et al. 2017. Lost in translation? Multi-metric macrobenthos indicators and bottom trawling. Ecological Indicators 82, 260-270. Josefson, A.B. 2008. DKI beregninger for danske lavvandede og lukkede områder. Rapport til BLST juni 2008. Josefson, A.B., Blomqvist, M., Hansen, J.L.S., Rosenberg, R. & Rygg, B. 2009. Assessment of marine quality change in gradients of disturbance: Comparison of different Scandinavian multi-metric indices. Marine Pollution Bulletin 58, 1263-1277.

Member States: ZKI

Indicator description	Indicator name	zoobenthos community indicator (ZKI)
	Indicator description	The benthic invertebrate index ZKI is based on the Pearson-Rosenberg model (Pearson and Rosenberg, 1978) of the community succession at a gradient of organic enrichment. ZKI is a biomass-based index, because relative biomasses relate better to nutrient enrichment than relative abundances in the Eastern Baltic Sea region. The ZKI index also

<table>
<tr><td rowspan="9"></td><td></td><td colspan="12">evaluates the number of species at stations and compensates this number for the salinity gradient. The compensation term is based on waterbody-specific maximum values for species richness calculated from the entire content of national database. The values of the ZKI index vary between 0 and 1, higher values representing healthier communities. The sensitivity of species was assessed based on the comparison of the historical data (when no symptoms of system-wide human-caused eutrophication were detectable) with the modern data from the Estonian marine waters.</td></tr>
<tr><td>Type of indicator</td><td colspan="4">☐ Model</td><td colspan="4">X Empirical-based</td><td colspan="4">☐ Pressure</td></tr>
<tr><td>Pressure assessed</td><td colspan="12">Nutrient enrichment (marine eutrophication)</td></tr>
<tr><td>Human activity</td><td colspan="12">Nutrient loading</td></tr>
<tr><td>MSFD criteria /descriptor</td><td colspan="12">This indicator was developed for WFD but it is also used in MSFD</td></tr>
<tr><td>How does the indicator relate to benthic biological diversity?</td><td colspan="12">Indicator predicts how benthic diversity (species richness) respond to nutrient enrichment along salinity gradient.</td></tr>
<tr><td>How does the indicator relate to benthic community structure and function?</td><td colspan="12">Indicator has two components: first part assesses the biomass share of benthic invertebrate species of different sensitivities to nutrient enrichment. The second part assesses the species richness of benthic invertebrate communities and compares this to a waterbody-specific maximum value. As such the indicator reflect the richness of community and shows if this community is in its pristine condition or not.</td></tr>
<tr><td>Indicator status</td><td colspan="4">☐ Under development</td><td colspan="4">X Applied for MSFD</td><td colspan="4">X Applied for other management, if so, for what: WFD</td></tr>
<tr><td>Regions with operational assessments</td><td colspan="12">Estonian marine waters</td></tr>
<tr><td rowspan="5">Input data</td><td>Biological data input (e.g. monitoring program, time series, sampling method)</td><td colspan="12">Estonian National Marine Monitoring Programme, assessment is based on (mostly annual) benthic (grab) sampling carried out in different Estonian waterbodies.</td></tr>
<tr><td>Targeted organisms</td><td colspan="3">X Infauna</td><td colspan="3">X Epifauna</td><td colspan="3">☐ Demersal fish</td><td colspan="3">☐ Other: ….</td></tr>
<tr><td>Environmental data input (e.g. empirical/modelled, source, time series)</td><td colspan="12">None</td></tr>
<tr><td>Pressure data input (e.g. time series, empirical/modelled, source, national/international)</td><td colspan="12">None</td></tr>
<tr><td>Data availability</td><td colspan="12">Open source data</td></tr>
<tr><td rowspan="4">Methodology</td><td>Parameters determined from biological data
(e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)</td><td colspan="12">Macroinvertebrate species biomasses assessed for 1 m2 area.</td></tr>
<tr><td>Parameters determined from pressure data
(e.g. total SAR, years not fished, trawling interval)</td><td colspan="12">None</td></tr>
<tr><td>Algorithm type
List of categorical information (Presence/Absence, …)
Direct measurements (counts, areas, concentrations, …)
Single or multimetric indicators using basic arithmetics
Indicators using multivariate and complex statistics
Indicators derived from modelling approaches
Indicators reporting on trends</td><td colspan="12">3</td></tr>
<tr><td>References for state-pressure relation</td><td colspan="12">Lauringson, V., Kotta, J., Kersen, P., Leisk, Ü., Orav-Kotta, H., Kotta, I. 2012. Use case of biomass-based benthic invertebrate index for brackish waters in connection to climate and eutrophication. Ecological Indicators. 12, 123–132.</td></tr>
</table>

	Uncertainty estimation methodology	Through replication		
	Coding availability (e.g. scripts, GitHub)	None		
	Threshold present	Yes		
	Threshold methodology	Empirical: Comparison to historical data. See: Kotta, J., Lauringson, V., Kaasik, A., Kotta, I. 2012. Defining the coastal water quality in Estonia based on benthic invertebrate communities. Estonian Journal of Ecology, 61, 86–105.		
Output	Output variable type	X Continuous	☐ Categorical	☐ Proportional
	Output variable range / classes	0...1		
	Output availability (e.g. report, website, reference)	Publicly available		
	Uncertainty handling (e.g. present confidence interval)	Through replication		
	Spatial resolution (e.g. grid cell size, habitat level)	Calculated for each waterbody		
	Temporal resolution	Mostly annual (depending on the water quality status of a waterbody)		
	Seabed habitat levels presented?	No		
More info	Indicator lead person	Jonne Kotta, jonne@sea.ee		
	Indicator data contact	Kristjan Herkül, kristjan.herkul@ut.ee		
	References / Literature / Project websites	Kotta, J., Lauringson, V., Kaasik, A., Kotta, I. 2012. Defining the coastal water quality in Estonia based on benthic invertebrate communities. Estonian Journal of Ecology, 61, 86–105. Lauringson, V., Kotta, J., Kersen, P., Leisk, Ü., Orav-Kotta, H., Kotta, I. 2012. Use case of biomass-based benthic invertebrate index for brackish waters in connection to climate and eutrophication. Ecological Indicators. 12, 123–132.		

Member States: BBI

Indicator description	Indicator name	Brackish-water Benthic Index (BBI)			
	Indicator description	Index based on community structure and species abundances, taking into account proportions of sensitive/tolerant species. Indicator value evaluated against type specific set thresholds. The indicator is used in the national WFD and MSFD assessment in Finland.			
	Type of indicator	☐ Model	X Empirical-based	☐ Pressure	
	Pressure assessed	Multiple pressures, including e.g. eutrophication			
	Human activity	Multiple activities			
	MSFD criteria /descriptor	D6C5, D5C8			
	How does the indicator relate to benthic biological diversity?	Based on monitoring data taking into account community structure (species richness and abundances).			
	How does the indicator relate to benthic community structure and function?	Community structure (species richness and abundance) is the basis of the indicator.			
	Indicator status	☐ Under development	x Applied for MSFD	x Applied for other management, if so, for what: WFD nationally in Finland	
	Regions with operational assessments	Coastal areas of Finland			
Input data	Biological data input (e.g. monitoring program, time series, sampling method)	Benthic fauna community data as inferred from grab samples.			
	Targeted organisms	x Infauna	☐ Epifauna	☐ Demersal fish	☐ Other: ….
	Environmental data input (e.g. empirical/modelled, source, time series)	Indicator only based on biological parameters. Water types pre-defined based on environmental data.			
	Pressure data input (e.g. time series, empirical/modelled, source, national/international)	No pressure data			

	Data availability	Available through the Finnish Pohje-database		
Methodology	Parameters determined from biological data (e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)	Species abundances form the basis of the index. The index is built of comparison of sample BQI and Shannon Weaver against type specific max values, species richness and abundance.		
	Parameters determined from pressure data (e.g. total SAR, years not fished, trawling interval)	none		
	Algorithm type List of categorical information (Presence/Absence, …) Direct measurements (counts, areas, concentrations, …) Single or multimetric indicators using basic arithmetics Indicators using multivariate and complex statistics Indicators derived from modelling approaches Indicators reporting on trends	Calculation of index: $\mathrm{BBI} = \frac{\left[\left(\frac{\mathrm{BQI}}{\mathrm{BQI_{max}}}\right) + \left(\frac{H'}{H'_{max}}\right)\right]}{2} * \frac{\left[\left(1 - \frac{1}{\mathrm{AB_{tot}}}\right) + \left(1 - \frac{1}{\mathrm{S}}\right)\right]}{2}$		
	References for state-pressure relation	Perus et al. 2009, Ambio Vol. 36, No. 2–3, 250-256		
	Uncertainty estimation methodology	Indicator value is the 20th percentile of bootstrapped index values. The 20th percentile is used as a precautionary or “fail-safe” approach (Carstensen 2007, Leonardsson et al. 2009) placing results of high uncertainty into lower status categories.		
	Coding availability (e.g. scripts, GitHub)	R script available upon request		
	Threshold present	Type specific thresholds set for the Finnish coastal areas		
	Threshold methodology	Thresholds are calculated based on a pragmatic statistical scheme. In short, this method sets the Good/Moderate threshold (GES threshold) at 0.6 times the 10th percentile of the top 10% of all index values within the baseline dataset.		
Output	Output variable type	x Continuous	x Categorical	☐ Proportional
	Output variable range / classes	0-1, translated to 5 WFD status categories based on type specific class boundary values		
	Output availability (e.g. report, website, reference)	Latest assessment (2011-2016) found in the Finnish “state of the marine environment” report: http://hdl.handle.net/10138/274086 (p. 172, in Finnish)		
	Uncertainty handling (e.g. present confidence interval)	Confidence not included in the output.		
	Spatial resolution (e.g. grid cell size, habitat level)	Assessment done at water body level.		
	Temporal resolution	Indicator is calculated for 6-year assessment periods. Temporal assessment through comparing assessment periods.		
	Seabed habitat levels presented?	Not defined, but confined to soft sediments (where quantitative grab sampling is possible)		
More info	Indicator lead person	Currently responsible: Henrik Nygård (henrik.nygard@syke.fi)		
	Indicator data contact	Henrik Nygård (henrik.nygard@syke.fi)		
	References / Literature / Project websites	Perus et al. 2009, Ambio Vol. 36, No. 2–3, 250-256		

Member States: TDI

Indicator description	Indicator name	TDI derived ecological status		
	Indicator description	Trawl Disturbance Index: link species sensitivity index (SI) resulting from biological traits scoring (mobility, position, feeding, size, fragility) to their abundance or biomass. Then thresholds are found in this index relationship to abrasion to determine the ecological status of a given habitat		
	Type of indicator	☒ Model	☐ Empirical-based	☐ Pressure
	Pressure assessed	Physical abrasion		
	Human activity	Demersal fisheries / bottom trawling		

	MSFD criteria /descriptor	D6 (D6C3, D6C4, D6C5)			
	How does the indicator relate to benthic biological diversity?	It does not			
	How does the indicator relate to benthic community structure and function?	The indicator relates to relative benthic biomass or abundance per sensibility categories, which has been shown to correlate with general community structure and function.			
	Indicator status	☐ Under development	☒ Applied for MSFD (in France)	☐ Applied for other management, if so, for what:	
	Regions with operational assessments	Southern North Sea, English Channel, French Mediterranean waters			
Input data	Biological data input (e.g. monitoring program, time series, sampling method)	benthic mega-epifauna bycatch data obtained from scientific trawl surveys or under-water videos. biological traits scoring (mobility, position, feeding, size,fragility) Foveau et al., 2020 https://doi.org/10.17882/59517			
	Targeted organisms	☐ Infauna	☒ Epifauna	☐ Demersal fish	☐ Other: ….
	Environmental data input (e.g. empirical/modelled, source, time series)	Habitat type: EUNIS level 4 or other (specified by the user)			
	Pressure data input (e.g. time series, empirical/modelled, source, national/international)	International Swept area ratio per year, VMS-based			
	Data availability	On demand (Sandrine.Vaz@ifremer.fr)			
Methodology	Parameters determined from biological data (e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)	Compute community sensitivity index following TDI formula (de Juan and Demestre 2012, https://doi.org/10.1016/j.ecolind.2011.11.020) at each location			
	Parameters determined from pressure data (e.g. total SAR, years not fished, trawling interval)	Annual average (surface) Swept Area Ratio or multi-annual (90th percentile) Swept Area Ratio per grid cell			
	Algorithm type List of categorical information (Presence/Absence, …) Direct measurements (counts, areas, concentrations, …) Single or multimetric indicators using basic arithmetics Indicators using multivariate and complex statistics Indicators derived from modelling approaches Indicators reporting on trends	4			
	References for state-pressure relation	Jac et al. (2020b). https://doi.org/10.1016/j.ecolind.2020.106617			
	Uncertainty estimation methodology	relative mean absolute model error (RMAE) by habitat			
	Coding availability (e.g. scripts, GitHub)	R script available on demand			
	Threshold present	Yes if detectable statistically			
	Threshold methodology	model the relationship of the TDI to abrasion using segmented regressions to detect thresholds			
Output	Output variable type	☐ Continuous	☒ Categorical	☐ Proportional	
	Output variable range / classes	conversion of habitat distribution and abrasion maps into ecological status categories: “GES”, “adverse effect”, “adverse effect or habitat loss”, “probably habitat loss”, “habitat loss”, “undetermined”			

	Output availability (e.g. report, website, reference)	Jac et al. (2020b). https://doi.org/10.1016/j.ecolind.2020.106617
	Uncertainty handling (e.g. present confidence interval)	the value of the RMAE was classified into very low uncertainty (0–0.1), low uncertainty (0.1–0.2), moderate uncertainty (0.2–0.5), high uncertainty (0.5–0.75) and very high uncertainty (0.75–1)
	Spatial resolution (e.g. grid cell size, habitat level)	Same as SAR grid resolution
	Temporal resolution	Once over the studied time period
	Seabed habitat levels presented?	EUNIS level 4 or other (specified by the user)
More info	Indicator lead person	Sandrine Vaz : sandrine.vaz@ifremer.fr
	Indicator data contact	Sandrine Vaz : sandrine.vaz@ifremer.fr
	References / Literature / Project websites	Jac et al. (2020a). https://doi.org/10.1016/j.ecolind.2020.106631 Jac et al. (2020b). https://doi.org/10.1016/j.ecolind.2020.106617 Please note that the segmented regression approach that enable the detection of threshold may be applied to any index which construction is independent from pressure layer.

Member States: mTDI

Indicator description	Indicator name	mTDI derived ecological status			
	Indicator description	Modified Trawl Disturbance Index: weight species sensitivity index (SI) resulting from biological traits scoring (mobility, position, feeding, size,fragility) with their relative abundance or biomass. Then thresholds are found in this index relationship to abrasion to determine the ecological status of a given habitat.			
	Type of indicator	☒ Model	☐ Empirical-based	☐ Pressure	
	Pressure assessed	Physical abrasion			
	Human activity	Demersal fisheries / bottom trawling			
	MSFD criteria /descriptor	D6 (D6C3, D6C4, D6C5)			
	How does the indicator relate to benthic biological diversity?	It does not			
	How does the indicator relate to benthic community structure and function?	The indicator relates to relative benthic biomass or abundance to species sensitivity according to biological traits, which has been shown to correlate with general community structure and function.			
	Indicator status	☐ Under development	☒ Applied for MSFD (in France)	☐ Applied for other management, if so, for what:	
	Regions with operational assessments	Southern North Sea, English Channel, French Mediterranean waters			
Input data	Biological data input (e.g. monitoring program, time series, sampling method)	benthic mega-epifauna bycatch data obtained from scientific trawl surveys or under-water videos. biological traits scoring (mobility, position, feeding, size, fragility) Foveau et al., 2020 https://doi.org/10.17882/59517			
	Targeted organisms	☐ Infauna	☒ Epifauna	☐ Demersal fish	☐ Other: ….
	Environmental data input (e.g. empirical/modelled, source, time series)	Habitat type: EUNIS level 4 or other (specified by the user)			
	Pressure data input (e.g. time series, empirical/modelled, source, national/international)	International Swept area ratio per year, VMS-based			
	Data availability	On demand (Sandrine.Vaz@ifremer.fr)			
Methodology	Parameters determined from biological data (e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)	Compute community sensitivity index following mTDI formula (Foveau et al. 2017 https://doi.org/10.1371/journal.pone.0184486) at each location			
	Parameters determined from pressure data (e.g. total SAR, years not fished,	Annual average (surface) Swept Area Ratio or multi-annual (90th percentile) Swept Area Ratio per grid cell			

	trawling interval)			
	Algorithm type List of categorical information (Presence/Absence, …) Direct measurements (counts, areas, concentrations, …) Single or multimetric indicators using basic arithmetics Indicators using multivariate and complex statistics Indicators derived from modelling approaches Indicators reporting on trends	4		
	References for state-pressure relation	Jac et al. (2020b). https://doi.org/10.1016/j.ecolind.2020.106617		
	Uncertainty estimation methodology	relative mean absolute model error (RMAE) by habitat		
	Coding availability (e.g. scripts, GitHub)	R script available on demand		
	Threshold present	Yes if detectable statistically		
	Threshold methodology	model the relationship of the mTDI to abrasion using segmented regressions to detect thresholds		
Output	Output variable type	☐ Continuous	☒ Categorical	☐ Proportional
	Output variable range / classes	conversion of habitat distribution and abrasion maps into ecological status categories: “GES”, “adverse effect”, “adverse effect or habitat loss”, “probably habitat loss”, “habitat loss”, “undetermined”		
	Output availability (e.g. report, website, reference)	Jac et al. (2020b). https://doi.org/10.1016/j.ecolind.2020.106617		
	Uncertainty handling (e.g. present confidence interval)	the value of the RMAE was classified into very low uncertainty (0–0.1), low uncertainty (0.1–0.2), moderate uncertainty (0.2–0.5), high uncertainty (0.5–0.75) and very high uncertainty (0.75–1)		
	Spatial resolution (e.g. grid cell size, habitat level)	Same as SAR grid resolution		
	Temporal resolution	Once over the studied time period		
	Seabed habitat levels presented?	EUNIS level 4 or other (specified by the user)		
More info	Indicator lead person	Sandrine Vaz : sandrine.vaz@ifremer.fr		
	Indicator data contact	Sandrine Vaz : sandrine.vaz@ifremer.fr		
	References / Literature / Project websites	Jac et al. (2020a). https://doi.org/10.1016/j.ecolind.2020.106631 Jac et al. (2020b). https://doi.org/10.1016/j.ecolind.2020.106617 Please note that the segmented regression approach that enable the detection of threshold may be applied to any index which construction is independent from pressure layer.		

Member States: pTDI

Indicator description	Indicator name	pTDI derived ecological status		
	Indicator description	Partial Trawl Disturbance Index: weight sensitive species sensitivity index (SI) resulting from biological traits scoring (mobility, position, feeding, size, fragility) with their relative abundance or biomass. Then thresholds are found in this index relationship to abrasion to determine the ecological status of a given habitat		
	Type of indicator	☒ Model	☐ Empirical-based	☐ Pressure
	Pressure assessed	Physical abrasion		
	Human activity	Demersal fisheries / bottom trawling		
	MSFD criteria /descriptor	D6 (D6C3, D6C4, D6C5)		
	How does the indicator relate to benthic biological diversity?	It does not		
	How does the indicator relate to benthic community structure and function?	The indicator relates relative benthic biomass or abundance to highly sensitive species according to biological traits, which has been shown to correlate with general community structure and function.		

	Indicator status	☐ Under development ☒ Applied for MSFD (in France) ☐ Applied for other management, if so, for what:
	Regions with operational assessments	Southern North Sea, English Channel, French Mediterranean waters
Input data	Biological data input (e.g. monitoring program, time series, sampling method)	benthic mega-epifauna bycatch data obtained from scientific trawl surveys or under-water videos. biological traits scoring (mobility, position, feeding, size, fragility) Foveau et al., 2020 https://doi.org/10.17882/59517
	Targeted organisms	☐ Infauna ☒ Epifauna ☐ Demersal fish ☐ Other: ….
	Environmental data input (e.g. empirical/modelled, source, time series)	Habitat type: EUNIS level 4 or other (specified by the user)
	Pressure data input (e.g. time series, empirical/modelled, source, national/international)	International Swept area ratio per year, VMS-based
	Data availability	On demand (Sandrine.Vaz@ifremer.fr)
Methodology	Parameters determined from biological data (e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)	Compute community sensitivity index following mTDI formula (Foveau et al. 2017 https://doi.org/10.1371/journal.pone.0184486) at each location
	Parameters determined from pressure data (e.g. total SAR, years not fished, trawling interval)	Annual average (surface) Swept Area Ratio or multi-annual (90th percentile) Swept Area Ratio per grid cell
	Algorithm type List of categorical information (Presence/Absence, …) Direct measurements (counts, areas, concentrations, …) Single or multimetric indicators using basic arithmetics Indicators using multivariate and complex statistics Indicators derived from modelling approaches Indicators reporting on trends	4
	References for state-pressure relation	Jac et al. (2020b). https://doi.org/10.1016/j.ecolind.2020.106617
	Uncertainty estimation methodology	relative mean absolute model error (RMAE) by habitat
	Coding availability (e.g. scripts, GitHub)	R script available on demand
	Threshold present	Yes if detectable statistically
	Threshold methodology	model the relationship of the pTDI to abrasion using segmented regressions to detect thresholds
Output	Output variable type	☐ Continuous ☒ Categorical ☐ Proportional
	Output variable range / classes	conversion of habitat distribution and abrasion maps into ecological status categories: “GES”, “adverse effect”, “adverse effect or habitat loss”, “probably habitat loss”, “habitat loss”, “undetermined”
	Output availability (e.g. report, website, reference)	Jac et al. (2020b). https://doi.org/10.1016/j.ecolind.2020.106617
	Uncertainty handling (e.g. present confidence interval)	the value of the RMAE was classified into very low uncertainty (0–0.1), low uncertainty (0.1–0.2), moderate uncertainty (0.2–0.5), high uncertainty (0.5–0.75) and very high uncertainty (0.75–1)
	Spatial resolution (e.g. grid cell size, habitat level)	Same as SAR grid resolution

	Temporal resolution	Once over the studied time period			
	Seabed habitat levels presented?	EUNIS level 4 or other (specified by the user)			
More info	Indicator lead person	Sandrine Vaz : sandrine.vaz@ifremer.fr			
	Indicator data contact	Sandrine Vaz : sandrine.vaz@ifremer.fr			
	References / Literature / Project websites	Jac et al. (2020a). https://doi.org/10.1016/j.ecolind.2020.106631 Jac et al. (2020b). https://doi.org/10.1016/j.ecolind.2020.106617 Please note that the segmented regression approach that enable the detection of threshold may be applied to any index which construction is independent from pressure layer.			

Member States: mT

Indicator description	Indicator name	mT derived ecological status			
	Indicator description	Modified vulnerability Index: non-linear combination of species biological traits scoring (mobility, position, feeding, size, fragility, protection status) with their relative abundance or biomass. Then thresholds are found in this index relationship to abrasion to determine the ecological status of a given habitat			
	Type of indicator	☒ Model	☐ Empirical-based	☐ Pressure	
	Pressure assessed	Physical abrasion			
	Human activity	Demersal fisheries / bottom trawling			
	MSFD criteria /descriptor	D6 (D6C3, D6C4, D6C5)			
	How does the indicator relate to benthic biological diversity?	It does not			
	How does the indicator relate to benthic community structure and function?	The indicator relates to relative benthic biomass or abundance to species sensitivity according to biological traits, which has been shown to correlate with general community structure and function.			
	Indicator status	☐ Under development	☒ Applied for MSFD (in France)	☐ Applied for other management, if so, for what:	
	Regions with operational assessments	Southern North Sea, English Channel, French Mediterranean waters			
Input data	Biological data input (e.g. monitoring program, time series, sampling method)	benthic mega-epifauna bycatch data obtained from scientific trawl surveys or under-water videos. biological traits scoring (mobility, position, feeding, size, fragility) Foveau et al., 2020 https://doi.org/10.17882/59517			
	Targeted organisms	☐ Infauna	☒ Epifauna	☐ Demersal fish	☐ Other: ….
	Environmental data input (e.g. empirical/modelled, source, time series)	Habitat type: EUNIS level 4 or other (specified by the user)			
	Pressure data input (e.g. time series, empirical/modelled, source, national/international)	International Swept area ratio per year, VMS-based			
	Data availability	On demand (Sandrine.Vaz@ifremer.fr)			
Methodology	Parameters determined from biological data (e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)	Compute community sensitivity index following mT formula (Jac et al., 2020a. https://doi.org/10.1016/j.ecolind.2020.106631 After Certain et al., 2015) at each location			
	Parameters determined from pressure data (e.g. total SAR, years not fished, trawling interval)	Annual average (surface) Swept Area Ratio or multi-annual (90th percentile) Swept Area Ratio per grid cell			
	Algorithm type List of categorical information (Presence/Absence, …) Direct measurements (counts, areas, concentrations, …) Single or multimetric indicators	4			

	using basic arithmetics Indicators using multivariate and complex statistics Indicators derived from modelling approaches Indicators reporting on trends			
	References for state-pressure relation	Jac et al. (2020b). https://doi.org/10.1016/j.ecolind.2020.106617		
	Uncertainty estimation methodology	relative mean absolute model error (RMAE) by habitat		
	Coding availability (e.g. scripts, GitHub)	R script available on demand		
	Threshold present	Yes if detectable statistically		
	Threshold methodology	model the relationship of the mT to abrasion using segmented regressions to detect thresholds		
Output	Output variable type	☐ Continuous	☒ Categorical	☐ Proportional
	Output variable range / classes	conversion of habitat distribution and abrasion maps into ecological status categories: "GES", "adverse effect", "adverse effect or habitat loss", "probably habitat loss", "habitat loss", "undetermined"		
	Output availability (e.g. report, website, reference)	Jac et al. (2020b). https://doi.org/10.1016/j.ecolind.2020.106617		
	Uncertainty handling (e.g. present confidence interval)	the value of the RMAE was classified into very low uncertainty (0–0.1), low uncertainty (0.1–0.2), moderate uncertainty (0.2–0.5), high uncertainty (0.5–0.75) and very high uncertainty (0.75–1)		
	Spatial resolution (e.g. grid cell size, habitat level)	Same as SAR grid resolution		
	Temporal resolution	Once over the studied time period		
	Seabed habitat levels presented?	EUNIS level 4 or other (specified by the user)		
More info	Indicator lead person	Sandrine Vaz : sandrine.vaz@ifremer.fr		
	Indicator data contact	Sandrine Vaz : sandrine.vaz@ifremer.fr		
	References / Literature / Project websites	Jac et al. (2020a). https://doi.org/10.1016/j.ecolind.2020.106631 Jac et al. (2020b). https://doi.org/10.1016/j.ecolind.2020.106617 Please note that the segmented regression approach that enable the detection of threshold may be applied to any index which construction is independent from pressure layer.		

Member States: MarBIT

Indicator description	Indicator name	Marine Biotic Index Tool (MarBIT)		
	Indicator description	Multivariate index for status assessment (specifically developed for the WFD and its normative definitions) with four individual indicators on: taxon diversity (taxonomic spread), abundance distribution, fraction of sensitive taxa, and fraction of tolerant taxa. The indicators are rated against a defined reference condition which is either based on a taxon reference list (most indicator) or on ecological-statistical properties (abundance distribution). Each indicator is measured based on a number of samples taken per area and habitat (there are reference lists for soft bottom, hard substrate and phytal fauna), then standardized on a 5-class scale and finally the median (preferred) or mean (alternatively) of the indicators is the final assessment result.		
	Type of indicator	☐ Model	x Empirical-based	x Pressure
	Pressure assessed	General degradation, eutrophication		
	Human activity	No specific activity addressed		
	MSFD criteria /descriptor	D6C5, D5C8		
	How does the indicator relate to benthic biological diversity?	One of the indicators in the MarBIT directly measures taxon diversity in terms of taxonomic spread (TSI = taxonomic spread index), also as a proxy for functional diversity		
	How does the indicator relate to benthic community structure and function?	Community structure is included via the abundance distribution indicator and function is part of the TSI indicator.		
	Indicator status	☐ Under development	x Applied for MSFD	x Applied for other management, if so, for what: WFD

	Regions with operational assessments	German Baltic Sea, rocky habitats in the German North Sea (Helgoland)			
Input data	Biological data input (e.g. monitoring program, time series, sampling method)	Data from regular yearly macrozoobenthos monitoring, sampling with Kautsky sampling frame in spring (infauna on soft bottoms) or summer (epifauna on phytal and hard substrate), one sample per station, minimum 20 samples per area			
	Targeted organisms	x Infauna	x Epifauna	☐ Demersal fish	☐ Other: ….
	Environmental data input (e.g. empirical/modelled, source, time series)	none			
	Pressure data input (e.g. time series, empirical/modelled, source, national/international)	None			
	Data availability	Only via the EPA in Germany that does the monitoring			
Methodology	Parameters determined from biological data (e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)	Taxon list, abundance per sample			
	Parameters determined from pressure data (e.g. total SAR, years not fished, trawling interval)	none			
	Algorithm type List of categorical information (Presence/Absence, …) Direct measurements (counts, areas, concentrations, …) Single or multimetric indicators using basic arithmetics Indicators using multivariate and complex statistics Indicators derived from modelling approaches Indicators reporting on trends	2, 3 (4 for abundance distribution (Lilliefors test with adaptation by Mason & Bell (1986)))			
	References for state-pressure relation	Carletti A, Heiskanen AS (2009): Water Framework Directive intercalibration technical report – Part 3: Coastal and Transitional waters. JRC Scientific and Technical Reports EUR 23838 EN/3 – 2009.			
	Uncertainty estimation methodology	none			
	Coding availability (e.g. scripts, GitHub)	Closed source software for Windows and Mac, available from MariLim aquatic research upon request (requires buying a license)			
	Threshold present	yes			
	Threshold methodology	Ecological properties of the benthos communities at various points on the degradation gradient			
Output	Output variable type	x Continuous	☐ Categorical	☐ Proportional	
	Output variable range / classes	EQR value between 0–1 in 5 equidistant classes: 0–0.2 = bad, 0.2–0.4 = poor, 0.4–0.6 = moderate, 0.6–0.8 = good, 0.8–1.0 = high			
	Output availability (e.g. report, website, reference)	Indicator report available here: https://marilim.de/informationen/wasserrahmenrichtlinie/marbit			
	Uncertainty handling (e.g. present confidence interval)	none			
	Spatial resolution (e.g. grid cell size, habitat level)	Habitat level, representative for a whole WFD water body			
	Temporal resolution	Once per year			
	Seabed habitat levels presented?	Generalized habitats: Soft bottoms, hard substrate, phytal			
More info	Indicator lead person	Torsten Berg, MariLim aquatic research			
	Indicator data contact	Torsten Berg, MariLim aquatic research			

	References / Literature / Project websites	https://marilim.de/informationen/wasserrahmenrichtlinie/marbit

Member States: BENTIX

Indicator description	Indicator name	BENTIX		
	Indicator description	The BENTIX index (Simboura & Zenetos, 2002) was developed for the purposes of the European Water Framework Directive 2000/60/EC (WFD) and for the assessment of the ecological quality status of benthic macroinvertebrates' communities. The index uses the relative contribution of tolerant and sensitive taxa, recombining the five ecological groups (GI-GV) described by Hily (1984), Glémarec (1986) and Grall and Glémarec (1997) and used in the AMBI index, weighting them according to the ratio of their occurrence in the benthic fauna. The metric was designed for the Mediterranean coastal waters and renders a five-step numerical scheme for the classification of benthic communities. The selection of the weight coefficients in the Bentix formula is not random and it is based on the realization that the probability of a benthic species picked up randomly, to be tolerant to stress is 3:1. This ratio is multiplied by 2 to create a scale ranging from 2 to 6. The 'sensitive' taxa group GS, including all sensitive (GI) and indifferent (GII) species is weighted by 6 and the 'tolerant' taxa group GT, including all tolerant (GIII), first (GV) and second order opportunistic species (GIV) are equally weighted by 2. The index has been successfully intercalibrated with other metrics within the Mediterranean geographical intercalibration group (MedGIG) and established as a national method in Greece and Cyprus for the classification of benthic communities under the WFD. It has been applied for the classification of coastal water bodies of Greece throughout the first (2012-2015) and second (2018-2023) monitoring cycle, and also proposed for the assessment of GES for soft bottom benthic habitats under the MSFD (2008/56/EC) and specifically for the descriptor seafloor integrity (D6). In order to include also structural components of benthic communities for the purposes of the MSFD a formula is under development combining the BENTIX index with diversity indices using specific reference values for different habitat types.		
	Type of indicator	☐ Model	X Empirical-based	☐ Pressure
	Pressure assessed	It has been successfully applied for the classification of coastal water bodies of Greece and tested over a wide variety of geographical areas and habitat types against multiple pressures such as eutrophication and organic pollution, physical alteration, mining residues, contaminants, etc.		
	Human activity	Aquaculture, discharges, oil spill, dredging, mining, dumping etc.		
	MSFD criteria /descriptor	D6C3, D6C5, D5C8		
	How does the indicator relate to benthic biological diversity?	For MSFD purposes a formula is under development combining BENTIX with diversity indices (species richness and Shannon) testing specific reference values for different habitat types. Note: Univariate indices such as Shannon diversity, species richness and multimetric indices encompassing those indices, have been criticized as being dependent on too many factors. Natural variability, habitat type, sample size, sampling methodology etc, influence diversity and species richness which are generally recommended to be used with caution as environmental classification tools (GIG 2013).		
	How does the indicator relate to benthic community structure and function?	BENTIX is a biotic index based on the concept of indicator groups and uses the relative contribution of tolerant and sensitive taxa weighting them accordingly to the ratio of their occurrence in the benthic fauna.		
	Indicator status	☐ Under development	X Applied for MSFD	X Applied for other management, if so, for what: Water Framework Directive, Environmental Impact Assessment studies, etc.
	Regions with operational assessments	Eastern Mediterranean		
Input data	Biological data input (e.g. monitoring program, time series, sampling method)	WFD monitoring programme in coastal waters / Greece 2012-2015; 2018-2023 MSFD monitoring programme/ Greece 2018-2023 WFD monitoring programme in coastal waters/Cyprus Sampling method: Two replicate samples (0.1m2) are collected at each station for the analysis of zoobenthos. Samples for fauna analysis are sieved on board through a 1 or 0.5 mm sieve and		

<table>
<tr><td rowspan="5"></td><td></td><td colspan="4">stored in 4 % formalin solution, stained with Rose Bengal. Samples are sorted in the lab and are grouped into the main benthic groups. Subsequently most of the specimens are identified to the species level and only when this is not possible (broken material) to a higher taxonomic level (genus or family). Organisms of the complete sample are identified.
Sampling devise: Van Veen grab, Ponar grab, box corer, etc.
Sampling surface 0.1 or 0.05 m2 (does not affect reliability of the index to a high extent).</td></tr>
<tr><td>Targeted organisms</td><td>X Infauna</td><td>X Epifauna</td><td>☐ Demersal fish</td><td>☐ Other: ….</td></tr>
<tr><td>Environmental data input (e.g. empirical/modelled, source, time series)</td><td colspan="4">Supporting objective and quantitative environmental data (i.e. depth, wave exposure, currents, temperature and salinity ranges, sediment composition etc).</td></tr>
<tr><td>Pressure data input (e.g. time series, empirical/modelled, source, national/international)</td><td colspan="4">Objective and quantitative data to validate state-pressure relationship, i.e. nutrients, oxygen concentrations, pelagic primary production, organic carbon, contaminants, etc</td></tr>
<tr><td>Data availability</td><td colspan="4">Depends on the data.</td></tr>
<tr><td rowspan="4">Methodology</td><td>Parameters determined from biological data
(e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)</td><td colspan="4">Taxonomic composition, disturbance sensitive taxa, taxa indicative of pollution, and abundance.
In development the multimetric BENTIX based on a formula combining the index with diversity indices using specific reference values for different habitat types.</td></tr>
<tr><td>Parameters determined from pressure data
(e.g. total SAR, years not fished, trawling interval)</td><td colspan="4">Depends on the case study</td></tr>
<tr><td>Algorithm type
List of categorical information (Presence/Absence, …)
Direct measurements (counts, areas, concentrations, …)
Single or multimetric indicators using basic arithmetics
Indicators using multivariate and complex statistics
Indicators derived from modelling approaches
Indicators reporting on trends</td><td colspan="4">4</td></tr>
<tr><td>References for state-pressure relation</td><td colspan="4">There are several references. Below the initial description of the index and some representative studies:
Simboura, N. & A. Zenetos, 2002. Benthic indicators to use in ecological quality classification of Mediterranean soft bottom marine ecosystems, including a new Biotic index. Mediterranean Marine Science, 3/2:77-111.
Simboura N, Panayotidis P, Papathanassiou E., 2005. A synthesis of the Biological Quality Elements for the implementation of the European Water Framework Directive in the Mediterranean Ecoregion: the case of Saronikos Gulf. Ecological Indicators, 5: 253-266. JIF: 984.
Simboura, N., E. Papathanassiou & D. Sakellariou, 2007. The use of a biotic index (Bentix) in assessing long term effects of dumping coarse metalliferous waste on soft bottom benthic communities. Ecological Indicators, 7(1): 164-180. JIF: 102
Simboura, N. & S. Reizopoulou 2007. A comparative approach of assessing ecological status in two coastal areas of Eastern Mediterranean. Ecological Indicators,7: 455-468. JIF: 102
Simboura, N. & S. Reizopoulou, 2008. An intercalibration of classification metrics of benthic macroinvertebrates in coastal and transitional ecosystems of the Eastern Mediterranean ecoregion (Greece). Marine Pollution Bulletin, 56:116-126. JIF: 630
Occhipinti Ambrogi A., Forni G., Silvestri C., Argyrou, M.,Jordana E., Mavric B.,Pinedo, S., Simboura , Gorazd Urbanic, G., 2009. The Mediterranean intercalibration exercise on soft-bottom benthic invertebrates with special emphasis on the Italian situation. Marine Ecology, 30(4), 495–504.
Simboura, N. & Argyrou, M. 2010. An insight into the function of benthic classification indices tested in Eastern Mediterranean coastal waters. Marine Pollution Bulletin, 60(5): 701-709.
Simboura N., Zenetos A., Pancucci-Papadopoulou M.A., Reizopoulou S., Streftaris N., 2012. Indicators for the sea-floor integrity of the Hellenic Seas under the European Marine Strategy</td></tr>
</table>

		Framework Directive: establishing the thresholds and standards for good environmental status. Mediterranean Marine Science 13/1: 140-152		
	Uncertainty estimation methodology	The software for the calculation of the index sets the limits of parameters under which the results are not within the confidence limits. These parameters are based on the lowest number of scores species and the lowest number of the species in the matrix that is needed to calculate the index.		
	Coding availability (e.g. scripts, GitHub)	Software available at https://cloudfs.hcmr.gr/index.php/s/518zEK0QmObxj2o		
	Threshold present	At the border of good to high status, the sensitive group accounts roughly for more than 60% or more than two-third of the fauna, while the tolerant group as a whole (tolerant plus opportunists) accounts for less than 40% or less than one-third of the fauna. At the border of good to moderate status, the sensitive group accounts roughly for less than 40% or less than one-third of the fauna, while the tolerant group as a whole (tolerant plus opportunists) accounts for more than 60% or more than two-third of the fauna. For purely muddy habitats with fine (silt and clay particles over 90%) where the benthic fauna is normally dominated by some tolerant species, a refinement of the H/G (4.5) and G/M (3.5) boundaries is recommended as H/G: 4 and G/M: 3.		
	Threshold methodology	Class boundary values were set by plotting the percentage of sensitive taxa and of tolerant taxa against the decreasing values of the Bentix index on the x-axis. The point were the two curves cross, corresponds to the central value of the Good class where the two groups of sensitive and tolerant species are each 50% of the fauna. The points at equal distances (0.5) in each side of the crossline represent the high-good boundary limit, and at the other side of the center the boundary between good/moderate. The index has been successfully intercalibrated with other metrics within the Mediterranean geographical intercalibration group (GIG, 2013), approved by ECOSTAT and established as a national method in Greece and Cyprus for the classification of benthic communities under the WFD. The boundaries of the method were delimited following the paired metric concept included in the IC boundary setting protocol (EC, 2003).		
Output	Output variable type	X Continuous	X Categorical	☐ Proportional
	Output variable range / classes	0 (bad status) to 1 (High status). 5-class scheme: bad, poor, moderate, good, and high ecological status		
	Output availability (e.g. report, website, reference)	The method has been applied mostly in the Mediterranean Sea but also in coastal waters and estuaries of European Atlantic waters, Black Sea, South America, India, China, Malaysia and Iranian estuaries.		
	Uncertainty handling (e.g. present confidence interval)	High levels of confidence level 90-95%		
	Spatial resolution (e.g. grid cell size, habitat level)	Depends on data availability		
	Temporal resolution	Depends on data availability		
	Seabed habitat levels presented?	The NEAT tool (Nested Environmental status Assessment Tool) is used to combine a high-level integration from different indicators and ecosystem components in each basic EUNIS level and BHT in a spatial scale (i.e. MRU).		
More info	Indicator lead person	Nomiki Simboura		
	Indicator data contact	Sofia Reizopoulou sreiz@hcmr.gr		
	References / Literature / Project websites	About 100 references in SCI journals Index website: https://www.hcmr.gr/en/the-bentix-index/		

Member States: IQI

Indicator description	Indicator name	Infaunal Quality Index (IQI)		
	Indicator description	The infaunal quality index incorporates taxonomic diversity, evenness and the AMBI index (proportions of sensitive and opportunistic taxa) to assess macrobenthic invertebrate samples. Metrics are compared to minimally disturbed reference conditions within to derive an Ecological Quality Ratio (EQR).		
	Type of indicator	☐ Model	X Empirical-based	☐ Pressure
	Pressure assessed	The IQI was developed as a general disturbance indicator. Pressures likely to cause a water body to fail its environmental objectives for benthic invertebrates are those related point source pollution containing substances such as metals and hydrocarbons, with additional impacts anticipated from certain physical disturbances. Disturbance from aggregate extraction, commercial fishing, and contaminants from the oil and gas		

		industry prevail in offshore habitats.
	Human activity	The IQI responds to a range of pressures including chemical contaminants, organic enrichment, particulate smothering and certain physical disturbances.
	MSFD criteria /descriptor	D1 - Biological Diversity & D6 – Seafloor Integrity
	How does the indicator relate to benthic biological diversity?	The IQI incorporates measures of taxonomic richness and evenness which are compared to values expected under minimal disturbance for a given combination of sediment and salinity characteristics to account for habitat driven variability. Elevated values corresponding to increased EQR values.
	How does the indicator relate to benthic community structure and function?	The IQI incorporates measures of structural evenness and disturbance sensitivity which are compared to values expected under minimal disturbance for a given combination of sediment and salinity characteristics to account for habitat driven variability. Elevated values corresponding to increased EQR values.
	Indicator status	☐ Under development ☐ Applied for MSFD x Applied for other management, if so, for what: Water Environment Regulations
	Regions with operational assessments	Progress against the UK target was assessed for the Celtic Seas and Greater North Sea Marine Strategy Framework Directive sub-regions and their constituent UK biogeographic marine regional seas set out in Charting Progress 2. The indicator was not used for the OSPAR Intermediate Assessment in 2017**.** The boundary of Water Framework Directive Good Ecological Status has been set through Intercalibration with other Member States of the North East Atlantic Geographical Intercalibration Group.
Input data	Biological data input (e.g. monitoring program, time series, sampling method)	Sample level benthic macroinvertebrate assemblage quantitative abundance data from intertidal core or subtidal grab sampling. Data standardized using truncation rules (see Phillips et al., 2014).
	Targeted organisms	X Infauna ☐ Epifauna ☐ Demersal fish ☐ Other: ….
	Environmental data input (e.g. empirical/modelled, source, time series)	Assessment requires salinity and sediment information. Information can be quantitative (average salinity and sediment granulometry data) or qualitative (descriptive salinity regime and sediment description using the Folk (1952) method or similar). Salinity and sediment information is applied to numerical models to estimate metric values for minimally disturbed reference conditions for specific conditions (habitat) of the benthic assemblage.
	Pressure data input (e.g. time series, empirical/modelled, source, national/international)	None.
	Data availability	UK Marine Strategy Assessments used data from WFD monitoring used within the Cycle 2 River Basin Management Plans for England, Northern Ireland, Scotland and Wales, Habitats Directive, Marine Conservation Zone (MCZ), Disposal Ground Monitoring (Northern Ireland), Fishing Intensity Study, Thames London Gateway, UK Oil and Gas industry and Marine Aggregate Levy Sustainability Fund (MALSF) Regional Environmental Characterization (REC) monitoring.
Methodology	Parameters determined from biological data (e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)	The indices used are the AZTI Marine Biotic Index (AMBI), Simpson's Evenness, and taxa number. IQI incorporates each metric as a ratio of the observed value to that expected under reference conditions.
	Parameters determined from pressure data (e.g. total SAR, years not fished, trawling interval)	None.
	Algorithm type List of categorical information (Presence/Absence, …) Direct measurements (counts, areas, concentrations, …) Single or multimetric indicators using basic arithmetics Indicators using multivariate	3. Single or multimetric indicators using basic arithmetics and 5. Indicators derived from modelling approaches - due to the modelled reference conditions.

	and complex statistics Indicators derived from modelling approaches Indicators reporting on trends	
	References for state-pressure relation	Derived from analysis of empirical data. Details in: Phillips, G. R., Miles, A. C., Prior, A., Martina, L. J., Brooks, L., & Anwar, A. (2014). Infaunal Quality Index: WFD classification scheme for marine benthic invertebrates. R&D Technical Report. Bristol: Environment Agency
	Uncertainty estimation methodology	The approach to determining confidence in the Regional Seas assessments is based on the extent to which the indicator and associated assessments fulfil certain criteria which impact the extent to which the assessment represents the available habitat relevant to the indicator, the extent to which the indicator can identify the overall effect of the relevant pressures, and the confidence in the assessment result being above/below the indicator target. The overall confidence assessment is based on the lowest extent to which the criteria are fulfilled (known as: "one-out-all-out").
	Coding availability (e.g. scripts, GitHub)	Uses Microsoft Excel Workbook
	Threshold present	The indicator quality threshold is defined as the boundary for WER/WFD Good Ecological Status (EQR ≥0.64) or Good Ecological Potential (for Heavily Modified Water Bodies [HMWBs]). For Marine Strategy Framework Directive Sub-Region and CP2 Regional Sea assessments, the overall indicator target is achieved where the assessed surface area (the total assessed surface area of inshore water bodies and offshore 10km x 10km assessment units) meeting the quality threshold achieves the quantity threshold of 85%.
	Threshold methodology	IQI quality thresholds were initially derived to meet the requirements of the different Ecological Status classes (Bad, Poor, Moderate, Good and High) as defined by the WFD normative definitions. Preliminary boundaries were set by assessing changes in proportions of functional sensitivity groups of the benthic invertebrate communities over a pressure gradient of organic enrichment (Phillips et al. 2014). Boundaries were subsequently modified through the Phase 1 Intercalibration process for the North East Atlantic Geographical Intercalibration Group (NEAGIG) (Phillips et al. 2014, Borja et al. 2007) and formalized in Commission Decision 2008/915/EC. Borja, A., Josefson, A. B., Miles, A., Muxika, I., Olsgard, F., Phillips, G., Rodríguez, J. G., & Rygg, B. (2007). An approach to the intercalibration of benthic ecological status assessment in the North Atlantic ecoregion, according to the European Water Framework Directive. Marine Pollution Bulletin, 55, 42–52.
Output	Output variable type	☐ Continuous ☐ Categorical x Proportional
	Output variable range / classes	Status — EQR Values High — ≥0.75 Good — 0.64 ≤ 0.75 Moderate — 0.44 ≤ 0.64 Poor — 0.24 ≤ 0.44 Bad — <0.24
	Output availability (e.g. report, website, reference)	IQI tool available at: https://www.wfduk.org/resources%20/coastal-and-transitional-waters-benthic-invertebrate-fauna UK Marine Strategy indicator assessment report available at: https://moat.cefas.co.uk/biodiversity-food-webs-and-marine-protected-areas/benthic-habitats/infaunal-quality-index/
	Uncertainty handling (e.g. present confidence interval)	Method for calculating statistical uncertainty for the IQI described in Phillips et al., 2014. UK Marine Strategy indicator assessment uncertainty is described in the UK Marine Strategy Assessment report: https://moat.cefas.co.uk/biodiversity-food-webs-and-marine-protected-areas/benthic-habitats/infaunal-quality-index/
	Spatial resolution (e.g. grid cell size, habitat level)	IQI assessment occurs at the sample level. Aggregation to larger scale assessment is dependent upon the specific assessment requirements.
	Temporal resolution	The IQI method provides a single snapshot. No temporal aspects should be inferred, this

		would come from temporal aspects of the data used for the assessment. The UK Marine Strategy 2018 indicator assessment report was based on the 2015 WFD classifications which used data from 2007 – 2012.
	Seabed habitat levels presented?	The IQI is limited to sedimentary habitats. Reference conditions are based on specific particle size and salinity information. Reference conditions are developed for intertidal core and subtidal grab methods. Reference conditions for sediment types under the Folk (1952) classification method are estimated to enable assessment where sediment information is limited to qualitative descriptions only.
More info	Indicator lead person	Graham Phillips
	Indicator data contact	UK Marine Strategy dataset metadata: https://www.dassh.ac.uk/doitool/data/1665 UK Marine Strategy dataset DOI: https://doi.org/10.17031/1665
	References / Literature / Project websites	Marine Online Assessment Tool: https://moat.cefas.co.uk/biodiversity-food-webs-and-marine-protected-areas/benthic-habitats/infaunal-quality-index/ Phillips, G. R., Miles, A. C., Prior, A., Martina, L. J., Brooks, L., & Anwar, A. (2014). Infaunal Quality Index: WFD classification scheme for marine benthic invertebrates. R&D Technical Report. Bristol: Environment Agency.

Member States: MEDOCC

Indicator description	Indicator name	MEDOCC index
	Indicator description	Index developed and applied to the soft-bottom communities to assess the response of the communities to enrichment gradient (organic matter in sediments)
	Type of indicator	☐ Model ☒ Empirical-based ☐ Pressure
	Pressure assessed	Organic matter enrichment
	Human activity	Eutrophication
	MSFD criteria /descriptor	soft-bottom macroinvertebrates
	How does the indicator relate to benthic biological diversity?	The index does not include biological diversity
	How does the indicator relate to benthic community structure and function?	Considering the number (%) of sensitive, indifferent, tolerant and opportunistic species along the organic matter gradient
	Indicator status	☐ Under development ☐ Applied for MSFD ☒ Applied for other management, if so, for what: WFD
	Regions with operational assessments	Mediterranean Sea
Input data	Biological data input (e.g. monitoring program, time series, sampling method)	Data from Monitoring program
	Targeted organisms	☒ Infauna ☐ Epifauna ☐ Demersal fish ☐ Other: ….
	Environmental data input (e.g. empirical/modelled, source, time series)	
	Pressure data input (e.g. time series, empirical/modelled, source, national/international)	Empirical data of organic matter in sediments in Catalonia and Balearic Islands
	Data availability	From 2002
Methodology	Parameters determined from biological data (e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)	Abundance of macroinvertebrates in four ecological groups (sensitivity classes)
	Parameters determined from pressure data (e.g. total SAR, years not fished, trawling interval)	Organic matter content in sediments
	Algorithm type	2. Direct measurements: organic matter in sediment

	List of categorical information (Presence/Absence, …) Direct measurements (counts, areas, concentrations, …) Single or multimetric indicators using basic arithmetics Indicators using multivariate and complex statistics Indicators derived from modelling approaches Indicators reporting on trends			
	References for state-pressure relation	Bibliographic information and empirical data. Pinedo, S., E. Jordana, M. Manzanera & E. Ballesteros. 2016. Using MEDOCC index to evaluate the Ecological Status of Catalan coastal waters (Northwestern Mediterranean Sea) over time and depths. In: Experiences for Ground, Coastal and Transitional Water Quality Monitoring: The EU Water Framework Directive Implementation in the Catalan River Basin District (Part II) (A. Munné et al., eds.). Hdb Env Chem 43: 201-226. Jordana, E., S. Pinedo & E. Ballesteros. 2016. Assessing the environmental quality in heavily modified transitional waters: the application of MEDOCC index in Ebre Delta bays. In: Experiences for Ground, Coastal and Transitional Water Quality Monitoring: The EU Water Framework Directive Implementation in the Catalan River Basin District (Part II) (A. Munné et al., eds.). Hdb Env Chem 43: 227-248. Pinedo, S., E. Jordana & E. Ballesteros. 2015. A critical analysis on the response of macroinvertebrate communities along disturbance gradients: description of MEDOCC (MEDiterranean OCCidental) index. Mar. Ecol.,36:141-154. Jordana, E., Pinedo, S. & E. Ballesteros. 2015. Macrobenthic assemblages, sediment characteristics and heavy metal concentrations in soft-bottom Ebre Delta bays (NW Mediterranean). Env. Mon. Assess., 187:71.		
	Uncertainty estimation methodology			
	Coding availability (e.g. scripts, GitHub)			
	Threshold present			
	Threshold methodology			
Output	Output variable type	☐ Continuous	☒ Categorical	☐ Proportional
	Output variable range / classes	0-6		
	Output availability (e.g. report, website, reference)	Pinedo, S., E. Jordana, M. Manzanera & E. Ballesteros. 2016. Using MEDOCC index to evaluate the Ecological Status of Catalan coastal waters (Northwestern Mediterranean Sea) over time and depths. In: Experiences for Ground, Coastal and Transitional Water Quality Monitoring: The EU Water Framework Directive Implementation in the Catalan River Basin District (Part II) (A. Munné et al., eds.). Hdb Env Chem 43: 201-226. Jordana, E., S. Pinedo & E. Ballesteros. 2016. Assessing the environmental quality in heavily modified transitional waters: the application of MEDOCC index in Ebre Delta bays. In: Experiences for Ground, Coastal and Transitional Water Quality Monitoring: The EU Water Framework Directive Implementation in the Catalan River Basin District (Part II) (A. Munné et al., eds.). Hdb Env Chem 43: 227-248. Pinedo, S., E. Jordana & E. Ballesteros. 2015. A critical analysis on the response of macroinvertebrate communities along disturbance gradients: description of MEDOCC (MEDiterranean OCCidental) index. Mar. Ecol.,36:141-154. Jordana, E., Pinedo, S. & E. Ballesteros. 2015. Macrobenthic assemblages, sediment characteristics and heavy metal concentrations in soft-bottom Ebre Delta bays (NW Mediterranean). Env. Mon. Assess., 187:71.		
	Uncertainty handling (e.g. present confidence interval)			
	Spatial resolution (e.g. grid cell size, habitat level)	Sample level		
	Temporal resolution			
	Seabed habitat levels presented?	No		
More info	Indicator lead person	Susana Pinedo		
	Indicator data contact	pinedo@ceab.csic.es		

	References / Literature / Project websites	

General: GPBI

Indicator description	Indicator name	General-purpose biotic index (GPBI)			
	Indicator description	GPBI is based on the assumption that as a site becomes impacted by a pressure, the most sensitive species are the first to disappear, and that stronger impacts lead to more important losses. Thus, it explicitly uses the within-species loss of individuals in the tested station in comparison to one or several reference stations as the basis of ecological status assessment.			
	Type of indicator	☐ Model	X Empirical-based	☐ Pressure	
	Pressure assessed	A biotic indices capable of detecting anthropogenic impact without having preliminary knowledge of the occurring pressures.			
	Human activity	Any human activity evaluation, tested for (1) maerl extraction in the northern Bay of Biscay, (2–3) dredge disposal and trawling in the North Sea, and (4) hypoxic events at the seafloor in the Gullmarfjord.			
	MSFD criteria /descriptor	Descriptor 1, 6			
	How does the indicator relate to benthic biological diversity?	The indicator measure the loss in species diversity.			
	How does the indicator relate to benthic community structure and function?	The benthic community structure is assessed based on loss in species within the community. Function is not assessed.			
	Indicator status	X Under development	☐ Applied for MSFD	☐ Applied for other management, if so, for what:	
	Regions with operational assessments	4 cases study areas			
Input data	Biological data input (e.g. monitoring program, time series, sampling method)	Any biological data set consisting of impact-control or assessment-reference monitoring data.			
	Targeted organisms	X Infauna	☐ Epifauna	☐ Demersal fish	☐ Other: ….
	Environmental data input (e.g. empirical/modelled, source, time series)	No, but dataset needs to be collected under the same environmental condition (e.g. habitat) or dataset and evaluation need to be split in groups reflecting the same environmental conditions.			
	Pressure data input (e.g. time series, empirical/modelled, source, national/international)	Any pressure data (continuous, categorical) associated with the benthic dataset			
	Data availability	Labrune et al. 2021. doi.org/10.3390/jmse9060654			
Methodology	Parameters determined from biological data (e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)	Species loss			
	Parameters determined from pressure data (e.g. total SAR, years not fished, trawling interval)	GPBI scores can be plotted against any pressure data variable.			
	Algorithm type List of categorical information (Presence/Absence, …) Direct measurements (counts, areas, concentrations, …) Single or multimetric indicators using basic arithmetics Indicators using multivariate and complex statistics Indicators derived from modelling approaches	4. Indicator using complex statistics			

	Indicators reporting on trends			
	References for state-pressure relation	A specific reference/control dataset is defined based on the type of evaluation (human activity impact or ecological state).		
	Uncertainty estimation methodology	?		
	Coding availability (e.g. scripts, GitHub)	R-script available in supplementary material Labrune et al., 2021		
	Threshold present	-		
	Threshold methodology	Signal detection theory was used to propose a sound good/moderate ecological quality status boundary		
Output	Output variable type	☐ Continuous	X Categorical	☐ Proportional
	Output variable range / classes	0-1		
	Output availability (e.g. report, website, reference)	-		
	Uncertainty handling (e.g. present confidence interval)	-		
	Spatial resolution (e.g. grid cell size, habitat level)	Depending on data availability, but score per habitat type within an assessment area		
	Temporal resolution	Depending on data availability, but normally once (determining annual average over the used time period)		
	Seabed habitat levels presented?	Any seabed habitats, as long as assessment occurs per habitat type following a control-impact (or assessment – reference) data approach.		
More info	Indicator lead person	Celine Labrune; celine.labrune@obs-banyuls.fr		
	Indicator data contact	Celine Labrune; celine.labrune@obs-banyuls.fr		
	References / Literature / Project websites	Labrune, C.; Gauthier, O.; Conde, A.; Grall, J.; Blomqvist,M.; Bernard, G.; Gallon, R.; Dannheim, J.; Van Hoey, G.; Grémare, A. A General-Purpose Biotic Index to Measure Changes in Benthic Habitat Quality across Several Pressure Gradients. J. Mar. Sci. Eng. 2021, 9, 654. https://doi.org/10.3390/ jmse9060654		

General: AMBI

Indicator description	Indicator name	AMBI (AZTI's Marine Biotic Index)			
	Indicator description	Index calculating the relative proportion of benthic macroinvertebrates corresponding to sensitive and opportunistic species (with five ecological groups)			
	Type of indicator	☐ Model	X Empirical-based	☐ Pressure	
	Pressure assessed	It has been tested in many human pressures coming from: multipressure, aquaculture, sewage discharges, eutrophication, physical alteration, chemical pollution, mining, ports, oil and gas extraction, aggregate extraction, climate change, etc. As for the MSFD terminology: organic matter, nutrients, physical loss, physical disturbance, extraction, biological disturbance, biological change, other substances.			
	Human activity	Aquaculture, discharges, oil and gas extraction, mining, aggregate extraction, dredging, fishing, etc.			
	MSFD criteria /descriptor	D5 (D5C8), D6 (D6C3)			
	How does the indicator relate to benthic biological diversity?	Through the changes in the relative composition of species, after their responses to disturbance and changes in the proportion of ecological groups (sensitive, indifferent, tolerant, opportunistic of second order and opportunistic of first order)			
	How does the indicator relate to benthic community structure and function?	The indicator relates to overall benthic community abundance or biomass (it can be calculated using any of them), and, as such, with the structure and function of the community			
	Indicator status	☐ Under development	X Applied for MSFD	X Applied for other management, if so, for what: Water Framework Directive, Environmental Impact Assessment, etc.	
	Regions with operational assessments	All European regional seas and worldwide			
Input data	Biological data input (e.g. monitoring program, time series, sampling method)	It can be used by any sampling method available, usually using grabs of 0.1 m2 and 1 mm sieving mesh, but also using corers or sampling quadrats in intertidal areas			
	Targeted organisms	X Infauna	☐ Epifauna	☐ Demersal fish	☐ Other: ….

	Environmental data input (e.g. empirical/modelled, source, time series)	It is advisable having at least grain size, organic matter and redox potential to interpret the results, but any other data is interesting: contaminants, current speed and direction, depth, etc.
	Pressure data input (e.g. time series, empirical/modelled, source, national/international)	It is advisable: distance to the pressure source, intensity of the pressure, etc.
	Data availability	There is not repository, depends on data from each monitoring network
Methodology	Parameters determined from biological data (e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)	Not sure about this question: AMBI is calculated either from abundance or biomass, each species is assigned to one ecological group and the proportion is calculated
	Parameters determined from pressure data (e.g. total SAR, years not fished, trawling interval)	Depends on each case
	Algorithm type List of categorical information (Presence/Absence, …) Direct measurements (counts, areas, concentrations, …) Single or multimetric indicators using basic arithmetics Indicators using multivariate and complex statistics Indicators derived from modelling approaches Indicators reporting on trends	3
	References for state-pressure relation	There are over 1100 references of this index, here we include the initial description and two review papers with metanalyses from the whole world: Borja, A., J. Franco, V. Pérez, 2000. A marine biotic index to establish the ecological quality of soft-bottom benthos within European estuarine and coastal environments. Marine Pollution Bulletin, 40: 1100-1114. Borja, Á., S. L. Marín, I. Muxika, L. Pino, J. G. Rodríguez, 2015. Is there a possibility of ranking benthic quality assessment indices to select the most responsive to different human pressures? Marine Pollution Bulletin, 97: 85-94. Borja, A., G. Chust, I. Muxika, 2019. Chapter Three - Forever young: The successful story of a marine biotic index. Advances in Marine Biology, 82: 93-127.
	Uncertainty estimation methodology	No uncertainty estimated (but it can be determined)
	Coding availability (e.g. scripts, GitHub)	Software available at https://ambi.azti.es/
	Threshold present	Undisturbed: <1.2, slightly disturbed: 1-2-3.3, moderately disturbed: 3.3-5, heavily disturbed: 5-6, extremely disturbed: >6
	Threshold methodology	modelling
Output	Output variable type	X Continuous ☐ Categorical ☐ Proportional
	Output variable range / classes	0 (undisturbed) to 7 (azoic), classes: undisturbed, slightly disturbed, moderately disturbed, heavily disturbed and extremely disturbed
	Output availability (e.g. report, website, reference)	Not sure about this question, there is not repository, there are >1100 references with results
	Uncertainty handling (e.g. present confidence interval)	
	Spatial resolution (e.g. grid cell size, habitat level)	Depends on the data availability
	Temporal resolution	Depends on the data availability
	Seabed habitat levels presented?	No
More info	Indicator lead person	Angel Borja: aborja@azti.es
	Indicator data contact	same

	References / Literature / Project websites	There are >1100 references in SCI journals https://ambi.azti.es/

General: M-AMBI

<table>
<tr><td rowspan="11">Indicator description</td><td>Indicator name</td><td colspan="4">M-AMBI (Multivariate AZTI’s Marine Biotic Index)</td></tr>
<tr><td>Indicator description</td><td colspan="4">Factor analysis calculating the vectorial distance of AMBI, richness and diversity to reference conditions of bad and high ecological status</td></tr>
<tr><td>Type of indicator</td><td>☐ Model</td><td>X Empirical-based</td><td colspan="2">☐ Pressure</td></tr>
<tr><td>Pressure assessed</td><td colspan="4">It has been tested in many human pressures coming from: multipressure, aquaculture, sewage discharges, eutrophication,
physical alteration, chemical pollution, mining, ports, oil and gas extraction, aggregate extraction, climate change, etc.
As for the MSFD terminology: organic matter, nutrients, physical loss, physical disturbance, extraction, biological disturbance, biological change, other substances.</td></tr>
<tr><td>Human activity</td><td colspan="4">Aquaculture, discharges, oil and gas extraction, mining, aggregate extraction, dredging, fishing, etc.</td></tr>
<tr><td>MSFD criteria /descriptor</td><td colspan="4">D5 (D5C8), D6 (D6C3)</td></tr>
<tr><td>How does the indicator relate to benthic biological diversity?</td><td colspan="4">Through the changes in richness, Shannon diversity, and the relative composition of species, after their responses to disturbance and changes in the proportion of ecological groups (sensitive, indifferent, tolerant, opportunistic of second order and opportunistic of first order)</td></tr>
<tr><td>How does the indicator relate to benthic community structure and function?</td><td colspan="4">The indicator relates to overall benthic community abundance or biomass (it can be calculated using any of them), and, as such, with the structure and function of the community</td></tr>
<tr><td>Indicator status</td><td>☐ Under development</td><td>X Applied for MSFD</td><td colspan="2">X Applied for other management, if so, for what: Water Framework Directive, Environmental Impact Assessment, etc.</td></tr>
<tr><td>Regions with operational assessments</td><td colspan="4">All European regional seas and worldwide</td></tr>
<tr><td colspan="5"></td></tr>
<tr><td rowspan="5">Input data</td><td>Biological data input (e.g. monitoring program, time series, sampling method)</td><td colspan="4">It can be used by any sampling method available, usually using grabs of 0.1 m2 and 1 mm sieving mesh, but also using corers or sampling quadrats in intertidal areas</td></tr>
<tr><td>Targeted organisms</td><td>X Infauna</td><td>☐ Epifauna</td><td>☐ Demersal fish</td><td>☐ Other: ….</td></tr>
<tr><td>Environmental data input (e.g. empirical/modelled, source, time series)</td><td colspan="4">It is advisable having at least grain size, organic matter and redox potential to interpret the results, but any other data is interesting: contaminants, current speed and direction, depth, etc.</td></tr>
<tr><td>Pressure data input (e.g. time series, empirical/modelled, source, national/international)</td><td colspan="4">It is advisable: distance to the pressure source, intensity of the pressure, etc.</td></tr>
<tr><td>Data availability</td><td colspan="4">There is not repository, depends on data from each monitoring network</td></tr>
<tr><td rowspan="3">Methodology</td><td>Parameters determined from biological data
(e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)</td><td colspan="4">From either abundance or biomass, AMBI, richness and Shannon diversity is calculated</td></tr>
<tr><td>Parameters determined from pressure data
(e.g. total SAR, years not fished, trawling interval)</td><td colspan="4">Depends on each case</td></tr>
<tr><td>Algorithm type
List of categorical information (Presence/Absence, …)
Direct measurements (counts, areas, concentrations, …)
Single or multimetric indicators using basic arithmetics
Indicators using multivariate and complex statistics</td><td colspan="4">4</td></tr>
</table>

	Indicators derived from modelling approaches Indicators reporting on trends			
	References for state-pressure relation	There are over 500 references of this index, here we include the initial description and two review papers with metanalyses from the whole world: Muxika, I., Á. Borja, J. Bald, 2007. Using historical data, expert judgement and multivariate analysis in assessing reference conditions and benthic ecological status, according to the European Water Framework Directive. Marine Pollution Bulletin, 55: 16-29. Borja, Á., S. L. Marín, I. Muxika, L. Pino, J. G. Rodríguez, 2015. Is there a possibility of ranking benthic quality assessment indices to select the most responsive to different human pressures? Marine Pollution Bulletin, 97: 85-94. Borja, A., G. Chust, I. Muxika, 2019. Chapter Three - Forever young: The successful story of a marine biotic index. Advances in Marine Biology, 82: 93-127.		
	Uncertainty estimation methodology	No uncertainty estimated (but it can be determined)		
	Coding availability (e.g. scripts, GitHub)	Software available at https://ambi.azti.es/		
	Threshold present	This depends on the intercalibration exercises, in the original publication, the boundaries are: Bad ecological status: <0.2; Poor: 0.2-0.39, Moderate: 0.39-0.53, Good: 0.53-0.77, High: >0.77		
	Threshold methodology	Modelling and intercalibrated in Europe: European Commission, 2018. Commission Decision (EU) 2018/229 of 12 February 2018 establishing, pursuant to Directive 2000/60/EC of the European Parliament and of the Council, the values of the Member State monitoring system classifications as a result of the intercalibration exercise and repealing Commission Decision 2013/480/EU. Official Journal of the European Communities, L47: 1-91.		
Output	Output variable type	X Continuous	☐ Categorical	☐ Proportional
	Output variable range / classes	0 (worst status) to 1 (best status), classes: bad, poor, moderate, good, high ecological status		
	Output availability (e.g. report, website, reference)	Not sure about this question, there is no repository, there are >500 references with results		
	Uncertainty handling (e.g. present confidence interval)			
	Spatial resolution (e.g. grid cell size, habitat level)	Depends on the data availability		
	Temporal resolution	Depends on the data availability		
	Seabed habitat levels presented?	No		
More info	Indicator lead person	Angel Borja: aborja@azti.es		
	Indicator data contact	same		
	References / Literature / Project websites	There are >500 references in SCI journals https://ambi.azti.es/		

General: NEAT tool

Indicator description	Indicator name	NEAT		
	Indicator description	NEAT: Nested Environmental status Assessment Tool – This is not an Indicator, but an integration method working with various indicators and descriptors, see references below for details and examples		
	Type of indicator	☐ Model METHOD: indicator integration method	☐ Empirical-based	☐ Pressure
	Pressure assessed	Variable depending on selection of indicators but can be used to assess D6 related pressures, eg disturbance & loss		
	Human activity	Variable depending on selection of indicators & descriptors		
	MSFD criteria /descriptor	Can be used to assess all or any suite of MSFD descriptors and specific criteria but is commonly used to integrate information from indices related to D1/D6/D2/D3		
	How does the indicator relate to benthic biological diversity?	By incorporating various benthic status indicators e.g. BENTIX, AMBI, M-AMBI etc		

	How does the indicator relate to benthic community structure and function?	By incorporating various benthic status indicators e.g. BENTIX, AMBI, M-AMBI etc			
	Indicator status	☐ Under development	X Applied for MSFD	X Applied for other management, if so, for what: Applied for WFD	
	Regions with operational assessments	Use in formal MSFD assessements in the Eastern Mediterranean: Greece, Cyprus, and various published assessments, SEE: Uusitalo et al. 2016: Front. Mar. Sci., 06 September 2016 | https://doi.org/10.3389/fmars.2016.00159, application in 10 cases in Europe overing the four European regional seas covering 6 descriptors (D1, D2, D3, D4, D5 and D6), Pavlidou et al. 2019: Ecol. Indicat., 96 (2019), pp. 336-350, 10.1016/j.ecolind.2018.09.007 application in Greece for D1-D2-D3-D4-D5-D6-D7-D8, Integrated assessment of marine environmental status under the MSFD, NEAT demonstrated early warning eutrophication detection, NEAT demonstrated temporal changes related to the management measures. Borja et al. 2021: Front. Mar. Sci., 13 April 2021 | https://doi.org/10.3389/fmars.2021.638232 application in Malta covering seven descriptors (D1, D3, D5, D6, D8, D9, and D10), Kazanidis et al. 2020 Ecological indicators https://doi.org/10.1016/j.ecolind.2020.106624, application in deep waters in 9 study areas in the North Atlantic focusing on five MSFD descriptors with 24 indicators (one for D1, one for D3, seven for D4, 13 for D6, two for D10) used in the assessments of the nine areas, their habitats and ecosystem components. Fraschetti et al. 2022: https://doi.org/10.1016/j.jenvman.2021.114370 , application with an extensive dataset across five Mediterranean ecoregions including 26 Marine Protected Areas (MPAs), their reference unprotected areas, and a no-trawl case study, and several ecosystem components: seagrass Posidonia, macroalgae, sea urchins and fish. Thresholds to define the GES were set by dedicated workshops and literature review.			
Input data	Biological data input (e.g. monitoring program, time series, sampling method)	Depending on the indicators used			
	Targeted organisms	X Infauna	X Epifauna	X Demersal fish	X Other: ….macrofauna, reptiles, birds, various habitats, and other ecosystem components
	Environmental data input (e.g. empirical/modelled, source, time series)	Depending on the indicators used, it could be any type			
	Pressure data input (e.g. time series, empirical/modelled, source, national/international)	Depending on the indicators used, it could be any type			
	Data availability	Depends on the country and publication			
Methodology	Parameters determined from biological data (e.g. Species richness, abundance, biomass community, Shannon Weaver, Simpson, sensitivity classes)	It could be all of these but depends on the indicator used (many types of indicators are being used)			
	Parameters determined from pressure data (e.g. total SAR, years not fished, trawling interval)	It could incorporate all of these types of pressure data			
	Algorithm type List of categorical information (Presence/Absence, …) Direct measurements (counts,	It could incorporate all of these types of data and indicators			

	areas, concentrations, …) Single or multimetric indicators using basic arithmetics Indicators using multivariate and complex statistics Indicators derived from modelling approaches Indicators reporting on trends	
	References for state-pressure relation	Depending on the indicators/parameters used
	Uncertainty estimation methodology	NEAT allows integrated assessments by assembling data from various response variables and their associated error over different spatial and temporal scales (Borja et al., 2019, 2021; Pavlidou et al., 2019; Kazanidis et al., 2020). It is based on a hierarchical, nested structure of Spatial Assessment Units (SAUs), i.e. the areas where the environmental status assessment takes place (Borja et al., 2016a; Uusitalo et al., 2016). For details see Berg et al 2019
	Coding availability (e.g. scripts, GitHub)	
	Threshold present	YES the method requires all the chosen indicators to have set thresholds
	Threshold methodology	Depending on the indicators used
Output	Output variable type	☐ Continuous / X Categorical / X Proportional
	Output variable range / classes	Indicators are transformed into values that range from 0 (worst status) to 1 (best status) using a continuous piecewise linear interpolation (Berg et al., 2019). On this scale, the value of 0.60, identified as threshold value, corresponds to the boundary between GES and non-GES. The indicator values are translated to standardized values with four boundaries among different conditions: high-good (value of 0.80), good-moderate (value of 0.60), moderate-poor (value of 0.40) and poor-bad (value of 0.20) (Borja et al 2016a)
	Output availability (e.g. report, website, reference)	Published papers and assessments, various tables and maps
	Uncertainty handling (e.g. present confidence interval)	Each NEAT value has an associated confidence level, which is the probability of being in a determinate class status (bad, poor, moderate, good, high). This probability is estimated using the standard error linked to the observed indicator value, which is assumed to represent the mean value of a normal distribution. The resulting assessment was obtained by performing a Monte-Carlo simulation technique with 1000 iterations and using the standard error to repeat the assessment multiple times with simulated values. In this way, each iteration led to different NEAT values, returning a quantitative estimate of confidence level for the original NEAT values, expressed as the percentage of values falling into the five different assessment classes (Borja et al. 2016b).
	Spatial resolution (e.g. grid cell size, habitat level)	Variable and a number of options available
	Temporal resolution	Represening current status of the chosen assessment period (based on different types of data includign trends)
	Seabed habitat levels presented?	Works with different assessment levels including at specific habitat levels
More info	Indicator lead person	TOOL lead contact persons Torsten Berg & Angel Borja
	Indicator data contact	Depending on the case
	References / Literature / Project websites	www.devotes-project.eu/neat Berg, T. C. Murray, J. Carstensen, J.H. Andersen 2019, NEAT-nested environmental status assessment tool (2019), Manual-Version 1.4 - Borja, A., M. Elliott, J.H. Andersen, T. Berg, J. Carstensen, B.S. Halpern, A.S. Heiskanen, S. Korpinen, J.S. Stewart Lowndes, G. Martin, N. Rodriguez-Ezpeleta. Overview of integrative assessment of marine systems: the ecosystem approach in practice. Front. Mar. Sci., 3 (2016a), 10.3389/fmars.2016.00020 Borja, A. M. Elliott, P.V.R. Snelgrove, M.C. Austen, T. Berg, S. Cochrane, J. Carsten, R. Danovaro, S. Greenstreet, A.S. Heiskanen, C.P. Lynam, M. Mea, A. Newton, J. Patricio, L. Uusitalo, M.C. Uyarra, C. Wilson 2016. Bridging the gap between policy and science in assessing the health status of marine ecosystems. Front. Mar. Sci., 3 (2016b), 10.3389/fmars.2016.00175 Pavlidou, A., N. Simboura, K. Pagou, G. Assimakopoulou, V. Gerakaris, I. Hatzianestis, P. Panayotidis, M. Pantazi, N. Papadopoulou, S. Reizopoulou, C. Smith, M. Triantaphyllou,

		M.C. Uyarra, I. Varkitzi, V. Vassilopoulou, C. Zeri, A. Borja. Using a holistic ecosystem-integrated approach to assess the environmental status of Saronikos Gulf, Eastern Mediterranean. Ecol. Indicat., 96 (2019), pp. 336-350, 10.1016/j.ecolind.2018.09.007 Borja, A., I. Menchaca, J.M. Garmendia, J. Franco, J. Larreta, Y. Sagarminaga, Y. Schembri, R. González, R. Antón, T. Micallef, S. Camilleri, O. Solaun, A. Uriarte, M.C. Uyarra. Big insights from a small country: the added value of integrated assessment in the marine environmental status evaluation of Malta Front. Mar. Sci., 8 (2021), p. 375, 10.3389/fmars.2021.638232 Kazanidis, G., C. Orejas, A. Borja, E. Kenchington, L.-A. Henry, O. Callery, M. Carreiro-Silva, H. Egilsdottir, E. Giacomello, A. Grehane, L. Menoth, T. Morato, S.A. Ragnarsson, J.L. Rueda, D. Stirling, T. Stratmann, D. van Oevelen, A. Palialexis, J.M. Roberts. Assessing the environmental status of selected North Atlantic deep-sea ecosystems. Ecol. Indicat., 119 (2020), p. 106624, 10.1016/j.ecolind.2020.106624 Borja, A., J.M. Garmendia, I. Menchaca, A. Uriarte, Y. Sagarmínaga. 2019. Yes, we can! Large-scale integrative assessment of European regional seas, using open access databases. Front. Mar. Sci., 6 (2019), p. 19, 10.3389/fmars.2019.00019 Fraschetti, S. E Fabbrizzi, La Tamburello, M. C. Uyarra, F. Micheli, E. Sala, C. Pipitone, F. Badalamenti, S. Bevilacqua, J. Boada, E. Cebrian, G. Ceccherelli, M. Chiantore, G. D'Anna, A. Di Franco, Si. Farina, S. Giakoumi, E. Gissi, I. Guala, P. Guidetti, S. Katsanevakis, E. Manea, M. Montefalcone, M. Sini, Va. Asnaghi, A. Calò, M. Di Lorenzo, J. Garrabou, L. Musco, A. Oprandi, G. Rilov, An. Borja, 2022. An integrated assessment of the Good Environmental Status of Mediterranean Marine Protected Areas, Journal of Environmental Management, 305, 114370, https://doi.org/10.1016/j.jenvman.2021.114370. Uusitalo, L., H. Blanchet, J.H. Andersen, O. Beauchard, T. Berg, S. Bianchelli, A. Cantafaro, J. Carstensen, L. Carugati, S. Cochrane, R. Danovaro, A.-S. Heiskanen, V. Karvinen, S. Moncheva, C. Murray, J.M. Neto, H. Nygård, M. Pantazi, N. Papadopoulou, N. Simboura, G. Srėbalienė, M.C. Uyarra, A. Borja 2016. Indicator-based assessment of marine biological diversity–lessons from 10 case studies across the European seas Front. Mar. Sci., 3 (2016), p. 159, 10.3389/fmars.2016.00159

Annex 2 Report from the Review Group to scope assessment methods to set thresholds and assess adverse effects on seabed habitats (RGBENTH2)

Participants: Sophie Mormede, Steven Degraer, Simon Jennings (chair)

Meeting: By correspondence July-September 2022

Request: Review group participants were asked to review two reports:

1. ICES Workshop on assessment methods to set thresholds and assess adverse effects on seabed habitats (WKBENTH2)

2. Technical service to produce a compilation of assessment methods and indicators that can be used to assess seabed habitats under D6/D1 for the MSFD.

And to assess whether,

a) The analyses were technically correct.

b) The scope and depth of the science were appropriate for the request.

c) The analyses contained the knowledge to answer the request for advice.

Background

ICES advised that the RGBENTH2 review of WKBENTH2 would be provided to WKBENTH3 as well as to the subsequent Advice Drafting Group. WKBENTH3 will have the task of evaluating proposed assessment methods and evaluating thresholds for assessing adverse effects on seabed habitats, using agreed upon criteria, methods, and analyses of their performance. Outcomes of WKBENTH3 will also contribute to the advice to DGENV.

Note on process

This report combines comments from the three reviewers: Sophie Mormede, Steven Degraer and Simon Jennings. The reviewers had different backgrounds and expertise and each reviewer conducted an individual review of the documents before meeting with other reviewers to agree responses and the structure of the combined report. Although the reviewers' comments focused on different aspects of the reports, the compiled comments were discussed and agreed collectively.

Two tables are used to summarise the reviewers' responses to questions 'a' to 'c' in the request, and these are followed by a section-by-section review of the WKBENTH2 report. An opening summary highlights the key messages from the review.

One reviewer also annotated the original WKBENTH2 report with smaller comments. The annotated report is available from ICES Secretariat.

Summary

Both the ICES Workshop on assessment methods to set thresholds and assess adverse effects on seabed habitats (WKBENTH2) and the ICES Technical Services team adopted a technically appropriate approach to fulfil their Terms of Reference and to provide inputs to WKBENTH3 and the ADG. The scale of their task was large given the amount of work on benthic indicators now being linked to the MSFD D6 (and noting that much of this work was not initiated for this purpose) and the range of interpretations of MSFD D6 processes that exist both nationally and internationally.

We note the substantial progress made with respect to seafloor integrity indicator and threshold evaluation. WKBENTH2 (feeding into WKBENTH3) is another milestone in the long history of development, evaluation, and selection of seafloor integrity indicators and thresholds. We encourage WKBENTH3 to focus on the steps needed to screen and condense available information and process and to draw strong and tractable conclusions that will underpin advice.

The scope and depth of the scientific treatment of the WKBENTH2 Terms of Reference and requirements for the Technical Service are largely appropriate. They do provide a common evaluation framework as requested. The main omission is that the focus on the requirements of Commission Decision (EU) 2017/848 is not strongly developed in the WKBENTH2 report and in the evaluation of indicators and thresholds. A stronger focus on 2017/848 would likely help ICES to provide clearer and more actionable information and advice on the value of proposed indicators and thresholds.

The work completed to date by WKBENTH2 and ICES Technical Services will go a long way towards guiding the expected tasks of WKBENTH3 and addressing the request for advice, although amendments to the summary tables developed by WKBENTH2 and used in the Technical Service task are recommended. Specifics are detailed in Table 1 and Table 2 and the body of this review. Parts of the WKBENTH2 report conflate indicators, methods and sometimes thresholds. Going forward, we suggest these should be explicitly split, defined, and criteria applied to each. An appropriate ‘taxonomy’ may be: 1- candidate indicator, 2- methodology to calculate the indicator, 3- application of the methodology, and 4- thresholds. There are also some inconsistencies and redundancies in the report, perhaps reflecting the drafting of sections by different subgroups. Specific examples are provided in the body of this review.

If an indicator and threshold can be defined, then many different methods may be used to assess the value of the indicator in relation to the threshold. Fisheries science provides a classic aquatic example, where one indicator that is broadly accepted internationally is biomass and thresholds are often set as a proportion of unimpacted (modelled) biomass. However, a very wide range of methods and models are developed, tested, and used to acquire data, and ultimately estimate the biomass. It would likely be easier for developers if a small suite of indicator(s) and thresholds was defined and the science was more strongly focused on the methods to estimate values and associated uncertainty for these indicators.

The reviewers were concerned about the large differences in scoring that were observed, and this implies that further work is needed to reduce ambiguity in the criteria and/ or to advance common understanding of the scoring process.

It is suggested to go beyond the weighting and scoring criteria presented to identify criteria that are essential for an indicator or threshold to meet the requirements of Commission Decision (EU) 2017/848 and the MSFD. If indicators or thresholds do not/ will not meet these criteria (now, or on some specified future time frame to be proposed/ decided by the group) then it is logical that they are not emphasised as potentially appropriate in the advice (at this time), even if they score highly on some criteria. If a proportion of the numerous indicators and thresholds are identified as inappropriate for use at this stage, this will ultimately contribute to stronger and more concise advice on (a) the suitability and shortcomings of both risk and state indicators for MSFD assessment purposes, and (b) on threshold values, at national and regional scales. More widely, with indicators and thresholds being both numerous and at varying stages of development, progress with any performance evaluation, intercomparison or intercalibration exercise would have to be very protracted, or simplified to the point of being uninformative, to accommodate all suggested indicators and thresholds irrespective of whether they meet essential criteria for the MSFD.

We appreciate, of course, that many groups developing or using specific indicators will strongly champion them, and that this can complicate selection exercises. For the purposes of moving the exercise forward a few indicators may be classified as mature (unconditional pass) and thus carried forward in the current advice and others, rather than being entirely

dropped from the process, may be highlighted as 'conditional' passes with the necessary conditions for further development being clearly specified (tabulated). The conditional passes are the indicators that would then be flagged as inappropriate for use in the context of the MSFD at this time.

Two main types of indicators are listed in the Technical Services document: empirically-based methods and model-based methods, the former making the bulk of the indicators. Commission Decision 2017/848 states that physical loss shall be understood as a permanent change to the seabed which has lasted or is expected to last for a period of two reporting cycles (12 years) or more. This leads to a requirement to understand the rate of recovery, which clearly favours modelling options, as does the scale at which the assessment has to be applied compared to data availability.

Model-based methods can integrate spatial and temporal processes of impact and recovery and be calculated at the population or community scale. Caveats may include that they can become very complex, need to be well reviewed, tested, reproducible, and that initial state needs to be defined. Indicators BH3 and PD could be tested against each other, including in relation to recovery time, and with varying assumptions, data quality etc. BH4 and Cumul seem to need less data, and could potentially be used more widely, and could be correlated with more complex methods such as BH3 and PD. The uncertainty surrounding the simpler methods might be no more important than the combined effects of the assumptions in the more complex methods.

Many empirically-based methods (e.g. BISI, HELCOM etc) directly measure the current state, or the current pressure, and then compare with threshold values. There is merit in having direct measures of state and of pressure, particularly if they are to be monitored consistently over time in the same area. The risks of using such indicators include applying thresholds from other areas which might not be suitable, scale of sampling and monitoring etc. Further, we would caution against ensemble methods (such as NEAT) when there are very few well tested models to include in the ensemble. For example, if one result is correct and the other is not, the average will always be wrong.

The groups are encouraged to undertake a detailed review of 2017/848, especially the D6 annex, to discuss and determine the extent to which all requirements of this Decision lead to other relevant criteria for the selection of indicators and thresholds, and thus determine the appropriateness of these indicators and thresholds to support MSFD.

Much research effort has been focused on the impacts of active bottom fisheries on seafloor integrity. Many seafloor integrity indicators hence relate to the impact of bottom fisheries, which is visible in the WKBENTH2 report. The remit of WKBENTH2 however was to assess indicators for seafloor integrity. The WKBENTH2 resolution mentions "adverse effects on seabed habitats" and "condition of seabed habitats and the adverse effects of key pressures". There is no specific mentioning of bottom fisheries in the resolution. The report occasionally reads as if fishing pressure was the main topic. While we recognise the emphasis in Commission Decision 2007/848 that Member States "focus their efforts on the main anthropogenic pressures affecting their waters" and should "have sufficient flexibility, under specified conditions, to focus on the predominant pressures and their environmental impacts on the different ecosystem elements in each region or subregion…." some subtle reconsideration of the text may provide some more balance. For example, on lines 423-423 "…to impacts from bottom contacting fishing gears" is likely obsolete because the subcriteria do not refer to fishing pressure; Table 3.1.1, Criterion 5 "This should include if the indicator is capable of including different gears with different impacts on habitats or species, if this is relevant for the indicator and its application" gives a (presumably unintended) focus solely on fishing pressure; and in the case of lines 1174-1175 "The resulting dataset consisted mostly of shrimps and lobster, with a few mollusc stocks" the term "stock" typically refers to the population size of commercial species, while GES should not be restricted/related to commercial species. Better to use the term "population size". This would contribute to the general appreciation of the report as going beyond fisheries-related aspects of GES.

A final point relates to the role of ICES science in identifying both indicators and thresholds. A complexity of this process, and one that will be challenging for WKBENTH3, the ADG and ICES in general, is the absence of a stronger steer on thresholds and appropriate precaution from policy and policy-stakeholder dialogue. A science group would usually consider the consequences of setting different thresholds or adopting different levels of precaution, rather than advise on what the specific thresholds or the level of precaution should be. There is not much policy steer to help the group, but the little that has been agreed and published (primarily in Commission Decision 2017/848) should be directly addressed as a priority, especially the specific statement that physical loss shall be understood as a permanent change to the seabed which has lasted, or is expected to last, for a period of two reporting cycles (12 years) or more.

Tabulated review

Two tables with identical rows and columns were used to assess the contributions of WKBENTH2 and the ICES Technical Service to the DGENV advisory request. These are presented independently because one is focused primarily on interpretation of the Commission Decision 2017/848 and because the reviewers were familiar with different groups of indicators.

Table 1. Review 1 of contributions of WKBENTH2 and the ICES Technical Service to the DGENV advisory request.

Review question Advice request	a) Is the analysis technically correct?	b) Are the scope and depth of the science appropriate for the request?	c) Does the analysis contain the knowledge to answer the request for advice?
(i) A detailed review of indicators used, or under development, by Regional Sea Conventions, Member States and ICES, for assessing the state/condition of seabed habitats suitable for MSFD assessments. The indicators considered can also include peer-reviewed indicators which have large-scale application.	A detailed review is provided as a technical service, based on a template developed by WKBENTH2. The specifics of the review are dependent on specialists with knowledge of the individual indicators, although we do not find errors in assessments for the small number of indicators with which we are familiar. The analysis in relation to the properties/ criteria considered is technically thorough.	The process in general has been thorough, and information requested/ collected on the templates is relatively complete. We suggest the final review that goes into the advice should include additional criteria that link the properties of indicators explicitly to Commission Decision (EU) 2017/848 as well as to the generic properties of good indicators.	In part. The detail could be enhanced and more useful to the recipients if it included detail of links to Commission Decision (EU) 2017/848.
(ii) Advise, using a set of agreed criteria, on a common framework to evaluate methods to assess benthic risk (model) and state (data) indicators, with respective threshold values.	The basis of the frameworks proposed is technically reasonable, except for the weighting and scoring processes where we suggest that indicators not meeting ‘critical’ criteria should be identified as unsuitable for MSFD support even if they score highly on other criteria. ‘Critical’ criteria will need to be identified. Note the suggestion that this may be handled by assigning an ‘unconditional’ and ‘conditional’ pass with additional (future) requirements clearly highlighted in the case of ‘conditional’ passes.	Appropriate (but incomplete) frameworks for evaluation were developed by WKBENTH2. As in (i)b we suggest the scientific evaluation should include an assessment of the properties of indicators and thresholds in relation to specific policy requirements. The treatment of uncertainty should be addressed beyond ‘methodology’ and ‘output’ (eg present confidence interval). This may require clarification from the requesters of the advice (and we note it does not appear in the ToR for the WKBENTH2) but 2017/848 states [threshold values should] “be set on the basis of the precautionary principle, reflecting the potential risks to the marine environment” so we interpret that uncertainty should also be assessed in this context when evaluating indicators and thresholds. In practice, this may mean a criterion that assesses whether the approaches adopted by	“Yes” in general terms, but improvements to the criteria as described in ii(a) and ii(b) would improve the rigour of evaluation.

Review question Advice request	a) Is the analysis technically correct?	b) Are the scope and depth of the science appropriate for the request?	c) Does the analysis contain the knowledge to answer the request for advice?
		the indicator developers enable an assessment that any given (calculated/ recorded) value of the indicator is consistent with avoiding a defined threshold (e.g. for loss as defined in 2017/848 in the extreme case) with a high probability. Note comments in the summary of this RGBENTH2 report on the distinction between indicators and methods.	
(iii) A targeted benthic data call (via TG Seabed), in order for ICES to evaluate the performance of selected (reviewed) benthic risk and state indicators, in relation to their ability to assess the state/condition of seabed habitats and adverse effects from specified pressures.	Since this relates to a request via TG seabed it was not clear if the identification of benthic datasets in Section 5.1.2 of RGBENTH2 was relevant. The datasets identified in WKBENTH2 for use in WKBENTH3 would be suitable for assessing the effects of trawling disturbance.	See (iii)a.	See (iii)a.
(iv) Advice on threshold values to assess the quality of seabed habitats.	The basis for this advice is available in the WKBENTH2 report. There is much general text on thresholds, and this provides a technically appropriate review of the general topic, but the development of the specific links to MSFD could be more focused, especially in the section where workshop participants focus on the suitability of the approaches covered in the review. At least one threshold is defined in Commission Decision (EU) 2017/848 (as mentioned on line 850 of the WKBENTH2 report). This important point is not further developed, but our interpretation is that this already defines a threshold for loss (“Physical loss shall be understood as a permanent change to the seabed which has lasted or is expected to last for a period of two reporting cycles (12 years) or more”) and the technical science question is	Much of the material needed to provide advice is available in the WKBENTH2 report, but it needs to be significantly filtered to draw out material relevant to the MSFD and request. Note the guidance in the DGENV request to “Advise on values (or ranges of values) for the indicators which would distinguish a habitat in good condition from the one which is adversely affected or lost (in general or by specific pressures)”. This also helps to guide the focus of the text. There is also a reference to the significance of loss to thresholds in the background to ToR ‘b’ for WKBENTH2.	Please note comments on precaution, uncertainty, and thresholds in (ii)b.

Review question Advice request	a) Is the analysis technically correct?	b) Are the scope and depth of the science appropriate for the request?	c) Does the analysis contain the knowledge to answer the request for advice?
	whether recovery time can be determined for any proposed indicator given the defined threshold (ie. Does the scientific basis of this indicator provide for estimation of recovery time in years from the present state and therefore a determination of whether "loss" has occurred, as defined in 2017/848). If the science basis of an indicator does not allow this, then can it logically meet the needs of MSFD reporting at all? Note also the relevance of the 2017/848 text on precaution appears to apply "Threshold values should also be set on the basis of the precautionary principle, reflecting the potential risks to the marine environment".		
(v) Advice on the suitability and shortcomings of both risk and state indicators for MSFD assessment purposes at national and regional scales.	The work that has been reported is technically appropriate to the extent we can judge, but is not sufficiently complete (in terms of criteria used and criticality of review) to make the full assessment as requested under (v)	Knowledge base could be strengthened, especially by assessing the evidence base related to the relationship between proposed indicators and thresholds and Commission Decision 2017/848.	"Yes" in general terms, but improvements to the criteria as described in ii(a) and ii(b) would improve the rigour of evaluation. Also applies to (ii).

Table 2 . Review 2 of contributions of WKBENTH2 and the ICES Technical Service to the DGENV advisory request.

Review question Advice request	a) Is the analysis technically correct?	b) Are the scope and depth of the science appropriate for the request?	c) Does the analysis contain the knowledge to answer the request for advice?
TOR A: Establish a set of criteria that can be used to evaluate the suitability of regional indicators/assessment	Yes – small comments. A precautionary margin should not be in the indicator, although the indicator	Yes – minor comment. The process may be biased against new methodologies, whereas they	Partly. This analysis looked at indicators rather than assessment methods.

Review question Advice request	a) Is the analysis technically correct?	b) Are the scope and depth of the science appropriate for the request?	c) Does the analysis contain the knowledge to answer the request for advice?
methods to assess adverse effects on seabed habitats for MSFD purposes	should capture uncertainty. The precautionary margin should be explicit and included in the threshold only (or it could be double counted). Need to consider uncertainty in both indicators and thresholds. Suggest not to duplicate criteria in indicators and thresholds (e.g. spatial extent and analytical vs expert). All core criteria (for both indicators and thresholds) should have a fail if any essential criterion scores 0 but criterion 12 probably should not. A fail may be conditional (e.g. could be fixed with future work), with emphasis in the future ICES advice on the 'unconditional' passes?	should be encouraged (but shown to be better or complimentary to existing methodologies prior to adoption).	The same indicators can be calculated using different methods or models. The split is not clear or explicit. Assessment methods require criteria too, such as peer review, agreed assumptions, tested sensitivities to assumptions, replicable, documented etc. Some of these are captured in the table of indicators.
TOR B: Review methods and criteria to set thresholds of adverse effects on seabed habitats, and suggest operational options that can be illustrated using worked examples	Some issues. 'Natural variation' assumes that there is enough comparable untouched habitat that has been surveyed to come up with values. It also assumes transferability between habitats (Yates et al. 2018), and also that the variability that arises from this assumption will somehow be smaller than the variability of a depleted state. It would have been useful to see the worked example used to calculate the values for impacted areas. There seems to be confusion in the worked example between extent and quality. We interpret the Worms analysis as treating 40%B0 as the lower limit (which is a maintain population size argument), not that a minimum of 40% of the population has to be above 80%B0. We suggest the two most promising thresholds relate to ecosystem state (at or above a specified threshold such as	Some issues. There is some discussion about the importance of connectivity and the indicators and thresholds should probably be calculated at the meta-population scale. Yet there is no discussion on the scale at which the analysis is to be carried out, and it is applied at the grid size in the worked example. Some thresholds will only be available for specific indicators. For example, population thresholds will only work for those indicators related to the populations while 'natural variation' (or some other level) will work for a wider range of indicators (e.g. including species richness).	Unsure. There are no operational options proposed, particularly with regards to the scale at which to calculate and to apply these thresholds. One of the worked examples uses the Pitcher method. This highlights again the need to investigate the assessment models as well as the indicators and thresholds. There are many assumptions within this method (including no stock recruit relationship and the level of vulnerability of benthic animals).

Review question Advice request	a) Is the analysis technically correct?	b) Are the scope and depth of the science appropriate for the request?	c) Does the analysis contain the knowledge to answer the request for advice?
	40% B_0) and recovery time (which is specified in the Commission Decision 2017/848 anyway). . The lack of knowledge of stock-recruit relationship should not be a hindrance but used as sensitivity. If the thresholds are 40% of B0, or above for example, this value will have near no influence.		
TOR C: Suggest quantitative and qualitative ways to evaluate and compare the suitability and performance of indicators/assessment methods	Difficult to tell. This is mostly a meta-analysis with little information so it is hard to tell if the comparison is like for like. The risk-based impact score is summarised over MFSD habitat, with no explanation if this is an average or other metric. It might be more transparent to report the proportion of each habitat which qualifies for each level of impact. It is also unclear which "extent" rule and "quality" rule have been applied.	Probably not. This section provides examples of indicators and different values at different pressures. But it does not apply the process of applying thresholds such as 95%ile of 'natural variation' and test if the 0.8 rule holds. For those examples the entire process should have been carried out: scoring the indicators, applying the thresholds and scoring the thresholds. None of the examples looked at the 12-year recovery period in Commission Decision 2017/848.	See other boxes. There does not seem to be clear guidance coming out of this section.
TOR D: Provide input to a draft compilation of regional indicators/assessment methods to set threshold and assess adverse effects on seabed habitats	The authors commented how people familiar with specific methods are required to be able to score them. The same applies to the content of those tables.	This exercise is a balance between too much detail and not enough information. These summaries seem adequate to provide an idea of what is available. It would be worth cross checking that all criteria are covered in the tables (e.g. to capture the likelihood of future data availability). Major assumptions would be an informative extra category.	It appears that only indicators used in Europe were considered. Other work could have been considered such as what is done in SPRFMO for example, limiting to well-developed methods. It is good to see spatial resolution and uncertainty covered as they are in this table. Consideration of these issues may be developed and added to the criteria for assessing indicators.

Section reviews of WKBENTH2 report

Section 2.

Some of the information in Section 2 does not seem to link to the ToR and request. This distracts the reader from the core purpose, business, and conclusions of WKBENTH2 and we suggest this information may be removed from the report or included as Annexes.

Lines 144-152: It is unclear why the report elaborates on the challenges related to the identification of biogeographically relevant subdivisions, which is not part of the request. The valid advice for regional coordination to sort this out goes beyond the remits of WKBENTH2.

Lines 205- “Lessons learnt from the Water Framework Directive intercalibration”, the relevance of this elaborate section is unclear. We suggest to either clarify its relevance (relative to the remits of WKBENTH2) at the start of the section or to consider moving the section to an Annex.

Section 3.

Previous WGECO and WGBIODIV work largely focused on generic properties of indicators, so to address the ToR it was necessary for WKBENTH2 to extend their approaches and to develop a criteria list suited to the specific requirements of the MSFD. The cross checking conducted by WKBENTH2 is a reassuring process in the context of completeness of the work. WKBENTH2 could have gone further in developing the specificity of the criteria to MSFD, especially in relation to the context provided by Commission Decision (EU) 2017/848 and what has already stated in the same Decision about appropriate scales of reporting (existing definition of units).

For criterion 2 in Table 3.1.1. there is emphasis on a monitoring time series to establish baselines and reference levels. Our reading of 2017/848 is that this would not be an essential prerequisite. While Article 4, 1h does state “be based on long time-series data, where available, to help determine the most appropriate value” this is not necessarily consistent with the preamble “marine ecosystems may recover, if deteriorated, to a state that reflects prevailing physiographic, geographic, climatic and biological conditions, rather than return to a specific state of the past.” The latter implies model-based estimates of (current) baseline state may be required, unless there are comparator areas with comparable “prevailing physiographic, geographic, climatic and biological conditions” and where pressures are low enough to provide confidence that the state can be treated as ‘baseline’.

For threshold values it was not clear why Article 4 of Commission Decision 2017/848 was not addressed directly as part of the WKBENTH2 work, especially in the cases where values will be established through Union, regional or subregional cooperation. There is some correspondence between 2017/848 and Table 3.2.1., and it is appropriate to identify additional criteria for evaluating thresholds (such as those previously considered by WGECO), but it would clarify the development of subsequent advice to work with those criteria mentioned in 2017/848 directly (and provide an operational interpretation of them).

For indicators, Commission Decision 2017/848 also suggests the need for a criterion to assess whether an indicator is responsive to a (known) main pressure (in the region where it is used), consistent with “As a result, the number of criteria that Member States need to monitor and assess should be reduced, applying a risk-based approach to those which are retained in order to allow Member States to focus their efforts on the main anthropogenic pressures affecting their waters” and “Member States should have sufficient flexibility, under specified conditions, to focus on the predominant pressures and their environmental impacts on the different ecosystem elements in each region or subregion in order to monitor and assess their marine waters in an efficient and effective manner and to facilitate prioritisation of actions to be taken to achieve good environmental status.”

Overall, and for the purposes of this request, it is suggested to also have a set of criteria for indicators and thresholds that are clearly linked to the requirements of 2017/848. For transparency it is likely best to list these explicitly (rather than melding them into more general criteria from WGECO and elsewhere).

There are three specifics we would highlight in 2017/848 that would also be usefully considered and treated as criteria.

First, in the Annex for D6, it is stated that “Physical loss shall be understood as a permanent change to the seabed which has lasted or is expected to last for a period of two reporting cycles (12 years) or more.” It follows that a valuable, and likely essential, property of an indicator is that it can be used to establish ‘permanent’ change, either through ongoing monitoring (“has lasted” ie shown to be below threshold for 12 years) or duration of recovery (“is expected to last” ie a quantitative prediction of recovery rate, ideally addressing uncertainty, shows the indicator will not meet the threshold after 12 years). Understanding of recovery is also emphasised in “Physical disturbance shall be understood as a change to the seabed from which it can recover if the activity causing the disturbance pressure ceases”.

Second, area criteria for the threshold are highlighted in WKTRADE2, but is it necessary to consider these for the indicator too? It is clear from 2017/848 that indicators need to enable reporting of areas lost/ disturbed/ affected in units of km^2, so spatial coverage and resolution of application are both relevant (eg. to what extent can sampling be extrapolated to appropriate scales, what is the resolution of modelling).

Third, the treatment of uncertainty should be more explicit. 2017/848 states [threshold values should] “be set on the basis of the precautionary principle, reflecting the potential risks to the marine environment” so we interpret that uncertainty should also be assessed in this context when evaluating indicators and thresholds. In practice, this may mean a criterion that assesses whether the approaches adopted by the indicator developers enable an assessment that any given (calculated/ recorded) value of the indicator is consistent with avoiding a defined threshold (e.g. for loss as defined in 2017/848 in the extreme case) with a high probability. General provision of a confidence interval for the indicator may not enable this, depending on what the confidence interval represents. Note to avoid the risk of double counting by having precaution in the estimate of the indicator values as well as being addressed in the threshold. It would be most transparent to associate the precaution with the threshold (to define the required probability of avoiding an unwanted state) rather than the indicator. This would also be consistent with the recognition that some precaution is expected as defined in 2017/848 Article 4 Para 1(e) [thresholds shall] “(e) be set on the basis of the precautionary principle, reflecting the potential risks to the marine environment.”

Having considered the need for more criteria in WKBENTH3, our next suggestion is to consider whether scoring and weighting alone will be sufficient to address the request for advice (the request relating to the suitability and shortcomings of indicators and not to the general review of indicators used). In the case of some criteria, and especially those linked to 2017/848, a pass/ fail approach would usefully be introduced for some criteria (perhaps attached to a timescale to reflect when a ‘fail’ may be converted to a ‘pass’ in the longer term eg. following further R&D). With this approach, and if an indicator or threshold fails on one of the key criteria, it would not be carried forward regardless of scores on other criteria. We have provided comments in the earlier parts of this report on how this may be handled as ‘unconditional’ and ‘conditional’ pass if necessary, where the ‘conditions’ to be addressed would be clearly listed.

Lines 406-408 “Threshold evaluation was addressed by WGECO in 2013 (ICES 2013b), and a second table, adapted from the indicator table (ICES 2012) was produced. These were not given any weightings at the time, and these were developed at this workshop”. It is unclear what exactly has been done during WKBENTH2 and how this has been done, relative to what was already available.

Lines 421-422 “Each criterion was evaluated against the WGECO/WGBIODIV table (ICES 2013a)”. It would clarify to add the conclusion from this comparison. Some bullets read as mere cross-check; other bullets read as an evaluation of the suitability of the criteria (for further uptake in the analysis).

Lines 496-503. Somewhat unclear and most likely incomplete statement. Suggestion to delete because of its low relevance to the exercise.

Lines 591-597. It is difficult to fully understand what has been done here, e.g. where does table 3.2.2 come from and what process was used link 3.2.2 "approaches" to 3.2.1 "evaluation criteria". Weassume this will be less of an issue for the WKBENTH2 participants that will also contribute to WKBENTH3, but it may need some further explanation particularly for potentially new participants. Reference to Section 4.3 could be made because this is where the missing piece of the puzzle isfound.

Line 592 “This was based on work carried out at WGECO in 2013 (ICES 2013b)”: Has there been any adaptation to what was reported by WGECO (2013)? If yes, it would be good to have that elaborated in the report (cf. to provide maximum clarity).

Lines 529-531 “WKBENTH1 considered it desirable for an indicator to integrate multiple pressures, while WGECO/WGBIODIV felt that “specificity” was the critical factor. Both positions have merit,…”. We agree with this point of view. It is suggested to cross-check this decision with the MSFD expectations, where “specificity” may be an explicit requirement.

Table 3.1.1, Criterion 8: If the answer to 6a is “B”, then (most likely?) criterion 8 should be scored <1. The scoring of criterion 8 seems to contradict the flexibility inherently adopted by criterion 6a.

Table 3.1.1, Criterion 9 “the indicator is easy to understand and communicate”: While (the concept of) the indicator needs to be easy to understand and communicate, this does not necessarily hold true for the algorithm (or method). Given the somewhat interchangeable use of “indicator” and “method” throughout the document, this may need to be elaborated in the table.

Lines 573-580 “…For this reason, the report only included the evaluation where experts in the indicators AND the criteria were included” : The need for an expert in the evaluation criteria to be applied seems to be problematic. This issue could potentially (partly) be solved by further elaborating the text explaining each of the criteria, so they become unambiguous to non-experts as well.

Table 3.2.1: Would there be any value in normalizing the scoring for each of the categories; this to equalize the different aspects of a good indicator (rather than to put more emphasis on those aspects for which more criteria have been defined)?

Section 4

The range of options for setting thresholds and their pros and cons are well covered. As the text states at line 848 one threshold, notably loss, has already been defined as a policy decision and states what is effectively a limit reference point. It would be extremely helpful to develop this further in the context of the WKBENTH3 activity and the drafting of advice.

Key questions related to threshold setting are whether the point at which recovery time will exceed 12 years (and therefore the point at which habitat is defined as ‘lost’) can be determined with an indicator (and associated methods) selected and how precaution is introduced (including ‘how much precaution’)? Some precaution is expected as defined in 2017/848 Article 4 Para 1(e) [thresholds shall] “(e) be set on the basis of the precautionary principle, reflecting the potential risks to the marine environment.” Other thresholds may be needed, but one for loss and one for loss plus precaution would seem to be a minimum set already defined by the MSFD for DC61.

If the seabed is not in the unimpacted state, and recovery time is less than 12 years, then the seabed will be classified as disturbed (DC62): “Physical disturbance shall be understood as a change to the seabed from which it can recover if the activity causing the disturbance pressure ceases”.

There is still a question ‘recovery time to what’ that does not appear to be explicitly addressed in MSFD or 2017/848 and would be needed to set the threshold, though expected options are assumed to be close to the unimpacted state given this is necessarily the seabed that would not be classified as ‘disturbed’ or ‘lost’.

We suggest the above line of reasoning is much more prominent in the next steps of this work. Such an approach may also help with intercomparison, collation, MSFD reporting and so on, because different indicators may be used regionally to determine locations of habitat loss and disturbance based on the same threshold for recovery time plus a defined uncertainty buffer.

If a threshold for loss plus precaution were identified (e.g. threshold indicator value associated with 95% probability that recovery time does not exceed 12 years), an important question in relation to the thresholds being discussed in Section 4 is whether the resulting value of the threshold would also be close to a target associated with other ‘desirable’ properties of the seabed (such as given in the example sections covering the extent of natural variation). If this were the case and shown with evidence, then less resources may need to be invested in proposing options for more complex targets and managing the complex debate about what they should be. Although WKBENTH2 give some emphasis to the extent of natural variation as a means of defining thresholds, this type of approach is very monitoring intensive at the scales considered for MSFD. An alternate option would be a threshold set on the basis of a defined probability of avoiding loss, coupled with case studies to understand the relationship between this threshold and the values of the indicator associated with natural variation.

The preceding comments relate to a 'quality' threshold of course, and do not address the setting of thresholds for 'extent'. There appear to be no policy decisions thus far to guide progress on extent. Types of evidence sought are likely to be similar to those that have been used, in some cases controversially, to define area targets for MPA coverage. The use of arguments about connectivity, as highlighted in WKBENTH2, provide a science base that could lead to a presentation of options, though applying these in general terms will be a significant challenge. An obvious scientific point is whether the distribution of a given percentage loss will determine its implications and how this should be handled (eg. patches as opposed to one contiguous area in the assessment unit).

Lines 642-643 "Thresholds are defined here as the state at which an ecosystem transitions from a good to a degraded state": Most likely "and/or extent" is to be added because the threshold(s) for both state and extent are to be considered. In general, "state" has sometimes been used in its widest sense (i.e. including also extent), sometimes in its more narrow sense (i.e. excluding extent) which complicates a correct understanding of the text.

Lines 653-665 and 788-791: The delivery of ecosystem services has not been listed as another concept to think about in relation to thresholds. There is an opportunity here to link to ecosystem services, with reference to how ecosystem services loss link to ecosystem function loss. The science is not quite there but efforts to link ecosystem services to ecosystem function are ongoing. May be useful to keep this in mind as the knowledge base continues to grow.

Lines 813-815: We could argue that this approach does define "good enough" relative to societal costs but not to the societal benefits delivered by well-functioning ecosystems. Here is where the ecosystem services approach ("how much do we need?") may come in.

Lines 819-824 "Sustainable use should therefore not be conflated with good environmental state": Point taken. You may however argue that maximum sustainable yield as used in a fisheries context is different from "how much do we need" embracing "all" ecosystem services.

Lines 857-859 on expert judgement in threshold setting… "Advantages of such an approach is the low demand for data, but this approach can be subjective, inconsistent and open to bias (Dorrough et al., 2020)". The process to get to the expert judgement (consensus) may considerably help its objectification. This may be elaborated here.

Lines 969-972 "Most time-series will also need to be detrended because long-term changes that are related to for example climate change will be causing long term increases or decreases. This was not considered a weakness of the approach, because it is a way of dealing with multiple pressures that are operating at different spatial scales". This statement hints that GES is to be related only to regionally manageable pressure. It hence ignores that anthropogenic effects playing at scales larger-than-regional scales that may also change the ecological state of the marine environment. While there indeed is some logic behind this statement, it would be helpful to assess if this follows the MSFD philosophy.

Lines 1139-1140. It would be informative to also have some figures on variation around the average range in natural variation in the text.

Section 5

In general, the preparation of the datasets described is appropriate to support the next steps of the work and a reasonably broad range of geographies and depths are identified (although not quite as diverse as the range of depths and pressures proposed, perhaps optimistically, in the request for advice to ICES). The datasets identified in WKBENTH2 for use in WKBENTH3 would be suitable for assessing the effects of trawling disturbance (or comparable forms of abrasion), rather than a wider range of pressures discussed in the request. Trawling is, however, a good example for the testing in most regions and subregions given emphasis in 2007/848 that Member States "focus their efforts on the main anthropogenic pressures affecting their waters" and should "have sufficient flexibility, under specified conditions, to focus on the predominant pressures and their environmental impacts on the different ecosystem elements in each region or subregion….”

Section 6

The template provided the basis of a good technical service, although we lacked expertise in many of the indicators reviewed in the technical service document. Note comments on treatment of uncertainty under the review of Section 3. The request does ask ICES to consider “how confidence and uncertainty are handled” without further discussion, but it is important whether this handling of uncertainty is appropriate to what the MSFD seeks to achieve (as in previous comments we link the consideration of uncertainty primarily to the threshold). The work that follows from the provision of this template does address the request, although there would be added utility from the WKBENTH2 and Technical Service work as a whole if criteria scorings for the indicators and thresholds (when proposed) for all the same indicators were included in Section 3 (rather than a subset).

The process for collation of indicators as described seems thorough, though we do not have the combined expertise to comment reliably on completeness.

Section 7

It is pragmatic and reasonable for WKBENTH2 to conclude that it is unlikely that sampling will ever provide a representative picture of seafloor status at the regional scale where required for D6, so it is necessary to have an emphasis on models and extrapolation and addressing the uncertainties associated with this. Otherwise, Section 7 summarises topics already addressed in the review of preceding sections.

Annex A. Examples of the distinction between indicator, method, application, and threshold.

Example A – biomass-based example:

- Indicator: The biomass of a sentinel species might be a candidate indicator. Criteria that apply to the indicator might include if it is suitable to represent the health of the benthic ecosystem.
- Method: Many methods could be used to calculate the biomass of this sentinel species. The method itself will have many criteria including is it peer reviewed, what are the assumptions, has sensitivities to assumptions been carried out, is it replicable, documented etc.
- Application: The application of the method deals with what data are used in the application and at what scale the calculation is made for example. Criteria might include if the data are representative of the underlying processes, or if the calculation is made at a fine-enough temporal and spatial scale to be meaningful.
- Thresholds: Finally, thresholds need to deal with the spatial and temporal scale of that calculation. For example, a threshold might be that the biomass of that sentinel species at the scale of the population stays above 50% of a reference biomass with a 10% risk of dropping below 20% of some defined reference biomass. Another threshold could be that the relative benthic status does not drop below 20% in more than 20% of the entire range of the indicator species, or that it does not drop below 20% in more than 50% of the fished area in each habitat type. Another threshold could be the biomass that recovers to unimpacted biomass if all impacts are stopped for 12 years.

Example B – Some questions that arise from the example in Figure 5.1 in the report

- Indicator: proportion of sentinel species. Is this indicator suitable over multiple habitat types for example? Does it capture degradation of habitat or ecosystem processes adequately?
- Method: what species are counted and ignored? At what identification level are these required and does that level of identification need to be constant over time or over different areas to make the indicator comparable? Are there different ways to compute that indicator? What is the scale of the calculation?
- Application: what sampling regime is adequate spatially (does a sampling regime of $0.0175m^2$ really capture biological processes) and can you compare between different sampling scales?

- Threshold: Over what scale is the threshold applied and how are the results of the calculation over that scale summarised? How is the variability captured in the threshold? Is the threshold value transferable between different habitats or even between different studies?

Annex B. How thresholds perform for selected indicators

This Annex provides a worked example of applying a 'natural variation' threshold, with reference to Figure 5.1 in the WKBENTH2 report (reproduced below).

An approximation of the 'natural variation' thresholds has been added to the figure. Green = 75%ile of very low pressure (we did not use the 95%ile as suggested in the text because it is difficult to figure out where it is in the figure) and red is 0.8 times that value (at a scale which starts at -0.2).

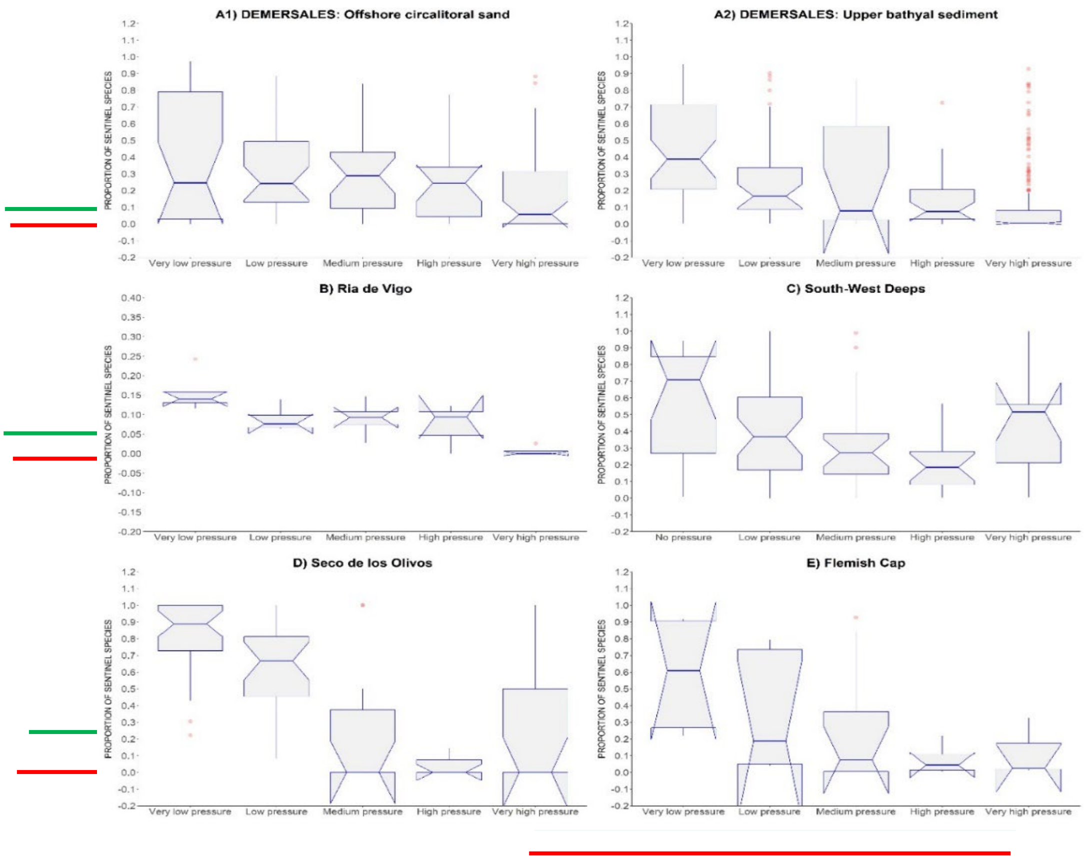
Panel A1 in Figure 5.1 shows no results below the threshold. Does that mean that even high pressure in this environment has limited effect or that the indicator is not suitable? Or is sampling inadequate to measure this change?

Panel D1 has all but low pressure below the threshold. Does that mean that only low pressure should be allowed? Is this environment more susceptible to trawling impacts than A1? Or is it that sampling captures change better?

In all other panels, results are partly above and partly below the red line. What constitutes a fail? Is it any point or a proportion of points or some other rule?

The area sampled varied between 0.0175m^2 and 0.04km^2. What effect have the sampled area and sample size (not reported) on the results and outcome? Can a sampling regime of 0.0175m^2 represent the biological processes? Should the thresholds be linked to the sampling regime?

Do we conclude that proportion of sentinel species is not a good indicator? Or just not a good indicator with regards to some specific habitat, level of habitat degradation, sampling type, some other reason, or a combination of all?



Reference

Yates, K.L.; Bouchet, P.J.; Caley, M.J.; Mengersen, K.; Randin, C.F.; Parnell, S.; Fielding, A.H.; Bamford, A.J.; Ban, S.; Barbosa, A.M.; Dormann, C.F.; Elith, J.; Embling, C.B.; Ervin, G.N.; Fisher, R.; Gould, S.; Graf, R.F.; Gregr, E.J.; Halpin, P.N.; Heikkinen, R.K.; Heinänen, S.; Jones, A.R.; Krishnakumar, P.K.; Lauria, V.; Lozano-Montes, H.; Mannocci, L.; Mellin, C.; Mesgaran, M.B.; Moreno-Amat, E.; Mormede, S.; Novaczek, E.; Oppel, S.; Ortuño Crespo, G.; Peterson, A.T.; Rapacciuolo, G.; Roberts, J.J.; Ross, R.E.; Scales, K.L.; Schoeman, D.; Snelgrove, P.; Sundblad, G.; Thuiller, W.; Torres, L.G.; Verbruggen, H.; Wang, L.; Wenger, S.; Whittingham, M.J.; Zharikov, Y.; Zurell, D.; Sequeira, A.M.M. (2018). Outstanding Challenges in the Transferability of Ecological Models. Trends in Ecology and Evolution 33, 790–802. https://doi.org/10.1016/j.tree.2018.08.001