

ICES WGBIODIV REPORT 2018

ECOSYSTEM PROCESSES AND DYNAMICS STEERING GROUP

ICES CM 2018/EPDSG:01

REF. SCICOM

Report of the Working Group on Biodiversity Science (WGBIODIV)

5–9 February 2018

ICES Headquarters, Copenhagen, Denmark



ICES
CIEM

International Council for
the Exploration of the Sea

Conseil International pour
l'Exploration de la Mer

International Council for the Exploration of the Sea Conseil International pour l'Exploration de la Mer

H. C. Andersens Boulevard 44-46
DK-1553 Copenhagen V
Denmark
Telephone (+45) 33 38 67 00
Telefax (+45) 33 93 42 15
www.ices.dk
info@ices.dk

Recommended format for purposes of citation:

ICES. 2018. Report of the Working Group on Biodiversity Science (WGBIODIV), 5-9 February 2018, ICES Headquarters, Copenhagen, Denmark. ICES CM 2018/EPDSG:01. 82 pp.

For permission to reproduce material from this publication, please apply to the General Secretary.

The document is a report of an Expert Group under the auspices of the International Council for the Exploration of the Sea and does not necessarily represent the views of the Council.

© 2018 International Council for the Exploration of the Sea

Contents

Executive summary	3
1 Administrative details	4
2 Terms of Reference.....	4
3 Summary of Work plan	5
4 Summary of Achievements of the WG during 3-year term	5
5 Final report on ToRs, workplan and Science Implementation Plan	6
5.1 Protocol on the development of theoretical concepts for biodiversity indicators (ToR1a, Deliverable 1)	6
5.1.1 Introduction.....	6
5.1.2 Background.....	7
5.1.3 A step-by-step approach to theory-based indicator conception.....	9
5.1.4 Case studies	15
5.1.5 Conclusions.....	19
5.1.6 References	19
5.2 Standardizing sampling design and analysis to obtain unbiased estimates of species diversity (ToR1b, Deliverable 2).....	24
5.2.1 Introduction.....	24
5.2.2 Species area relationship (SAR) and species accumulation curve (SAC).....	26
5.2.3 Case studies	28
5.2.4 References	49
5.3 Development of a Benthic Community Physical Disturbance Indicator	51
5.3.1 Introduction.....	51
5.3.2 Policy rationale.....	52
5.3.3 Scientific rationale.....	53
5.3.4 Constructing a traits-based indicator.....	55
5.3.5 Application of the indicator to two different case studies	57
5.3.6 Discussion.....	68
5.3.7 The future.....	70
5.3.8 References	70
6 Cooperation	72
7 Summary of Working Group self-evaluation and conclusions	72
Annex 1: Lists of participants.....	73
Annex 2: Recommendations.....	76
Annex 3: WGBIODIV draft resolution 2019–2021.....	77

Annex 4: WGBIODIV self-evaluation.....80

Executive summary

The ICES Working Group on Biodiversity Science (WGBIODIV), chaired by W. Nikolaus Probst, Germany, and Oscar Boss, the Netherlands, has held three annual meetings in 2016–2018.

WGBIODIV worked on theoretical concepts of diversity indicators along several lines. The quantification of changes in biodiversity is often considered central to the assessments of human impacts on marine ecosystems, but this quantification has been proven as an exceptionally difficult task. WGBIODIV developed a generic protocol on how indicators need to be underpinned conceptually to provide indicators of biodiversity with meaningful assessment targets. WGBIODIV also analysed the impacts of sampling design on estimates of species diversity. And finally, WGBIODIV developed a trait-based sensitivity indicator for benthic community in response to bottom trawling.

The generic protocol on indicator development identified seven steps that should be considered when developing indicators for environmental status assessments. Especially the development of a theoretical indicator framework is crucial, because it helps to develop meaningful assessment benchmarks. At the moment, many status indicators used for environmental status assessments by the regional seas conventions are missing this framework or using a non-quantitative theoretical background, which impairs the definition of quantitative assessment targets.

The influence of sampling design on estimates of diversity metrics was analysed in four different case studies: (i) comparing different methods for calculating species accumulation curves, (ii) analysing the influence of sampling size on North Sea fish species, (iii) analysing the impact of spatial resolution on plankton indicators and (iv) analysing the relationship between sampling effort and species number in soft bottom benthos. These case studies show that patterns of diversity strongly depend on the sampling scheme, which therefore requires careful consideration to provide the adequate data to feed into assessments.

WGBIODIV used traits-based data on benthic invertebrate communities to develop a community physical disturbance indicator. This indicator combines functional characteristics of benthic species, including their sensitivity to physical perturbations (i.e. their response through injury or death) and their recoverability (i.e. the self-sustainability of a population when damaged and/or its recolonization potential following mass mortality). The behaviour and performance of the indicator was examined using two independent case studies from the Dutch EEZ and the Bay of Biscay. Future investigations of WGBIODIV will focus on comparison of the distribution of indicator values between different geographical areas and, for some locations, over time. We will use the indicator to test hypotheses on the relationship between trawl effects and function of benthic communities.

1 Administrative details

<p>Working Group name</p> <p>Working Group on Biodiversity Science (WGBIODIV)</p> <p>Year of Appointment within current cycle</p> <p>2016</p> <p>Reporting year within current cycle (1, 2 or 3)</p> <p>3</p> <p>Chair(s)</p> <p>W. Nikolaus Probst, Germany</p> <p>Oscar Bos, the Netherlands</p> <p>Meeting dates and venues</p> <p>8–12 February 2016, San Sebastian, Spain (15 participants)</p> <p>6–10 February 2017, Venice, Italy (19 participants)</p> <p>11–12 February 2018, Copenhagen, Denmark (12 participants)</p>

2 Terms of Reference

ToR	Description	Background	Science Plan priorities addressed	Duration	Expected Deliverables
a	<p>Develop the use of biodiversity metrics (e.g. species richness and species evenness indices) to inform on the status of ecosystem components at the community level (fish, mammals, seabirds, plankton, epi-benthos, macro-algae) to support implementation of ecosystem-based management. This task encompasses:</p> <p>1a. Establish a sound theoretical basis relating variation in biodiversity metric values to changes in anthropogenic pressure on marine communities (e.g. incorporating components of community size and trophic structure into the derivation of biodiversity metrics, taking account of linkage to habitat types and consideration of spatial pattern).</p> <p>Update: ToR1a may require further work beyond next years' meeting and may extend into in the next term (2018–2020), as the</p>	<p>Initiatives to revise the EC Decision of 2010 suggest that metrics for the ecosystem level of biodiversity may simply not be possible given the current level of scientific knowledge. Instead metrics at community level may be achievable, and indeed community level metrics represent the logical progression from species level and habitat level in that communities represent the collection of species that occupy a habitat. In applying criteria to assess the performance of different community-level metrics, metrics of species diversity have routinely performed below par. A major</p>	<p>1, 5, 9, 11, 12, 13,16,18,</p>	<p>3 years</p>	<p>1. Protocol on the development of theoretical concepts of biodiversity indicators (2016/2017).</p> <p>2. Combined analysis and review on impacts of sampling size on performance of biodiversity metrics (2016–2018).</p> <p>3. Analysis on aggregating biodiversity indicators at different levels (species</p>

development of indicator concepts is time consuming.

1b. Explore the issue of sampling size dependence to derive a robust protocol for calculating biodiversity metrics so that their sensitivity to underlying drivers is maximized, and the ‘noise’ associated with sampling effects is minimized (e.g. procedures for sample aggregation, modeling of individual species distribution to derive point-diversity estimates).

1c. Assess the “ecosystem level” assessment of biodiversity by considering how community-level biodiversity metrics might be aggregated across communities (e.g. integrated ecosystem assessments of biodiversity).

Update: ToR1c may not be addressed during the 2016–2018 term as the development of trait-based indicators will not be completed until 2018.

1d. Apply the WGBIODIV quality criteria to assess the performance of state indicators to assess the performance of any biodiversity indicators proposed and developed by WGBIODIV to show whether previous weaknesses in such metrics have been addressed.

Update: ToR1d may have to be addressed in the next term (2018–2020) as the development of the new WGBIODIV biodiversity indicators may not be finalised in 2018.

shortcoming in their performance has been the lack of a sound and well understood theoretical basis to explain the relationship between pressure and state. Without this understanding, it has always been assumed that it would be difficult to formulate sound reliable scientific advice to support management based on observed variation in species diversity indicators. Consequently the community level indicators that have been used to support EAM initiatives, such as the OSPAR EcoQO pilot study and currently to fulfil the indicator 1.7.1 role for the MSFD focus on size based indicators such as the large fish indicator. Given the species diversity indicators would appear to be the most obvious candidates for metrics to fulfil the community-level indicator role in D1, the maintenance of biological diversity, the time is clearly ripe for the theoretical shortcomings in these indicators to be addressed so that they can be used to monitor change in biodiversity within marine communities.

group, community, ecosystem) (2017/2018).

4. Quality assessment of investigated biodiversity indicators according to WGBIODIV criteria (2018).

5. One or more operational indicators to assess biodiversity at the community and eventually the ecosystem level (2018).

3 Summary of Work plan

Year 1	Develop theoretical background for several indicators of diversity; establish protocol on indicator development
Year 2	Calculate biodiversity metrics using reference data, provide overview and comparison of outcomes
Year 3	Evaluate biodiversity indicators according to WGBIODIV indicator quality criteria

4 Summary of Achievements of the WG during 3-year term

In 2015, WGBIODIV chaired a theme session at the ICES Annual Science Conference on measuring and assessing biodiversity. The theme session hosted 18 oral presentations and three posters.

Members of WGBIODIV published two papers during the 2016–2018 reporting cycle:

- Probst, W. N., Lynam, C. P. 2016. Aggregated assessment results depend on aggregation method and framework structure - a case study within the European Marine Strategy Framework Directive. *Ecological Indicators*, 61: 871–881.
- Rambo, H., Stelzenmueller, V., Greenstreet, S.P.R., Moellmann, C. 2017. Mapping fish community biodiversity for European marine policy requirements. *ICES Journal of Marine Sciences*, 74: 2223–2238.

Deliverable 3 could not be achieved as only one indicator on benthic communities became developed. The aggregation of indicators of different ecosystem components was thus not possible. To date it remains open when and how WGBIODIV will be able to address this deliverable.

Deliverable 4 could not be achieved, as none of the envisioned indicators could be developed so far as to make it operational. Thus, an evaluation according to the WGBIODIV indicator quality criteria was not possible, but may be achievable in the next three-year working cycle.

5 Final report on ToRs, workplan and Science Implementation Plan

5.1 Protocol on the development of theoretical concepts for biodiversity indicators (ToR1a, Deliverable 1)

Deliverable 1: Protocol on the development of theoretical concepts of biodiversity indicators (2016/2017)

5.1.1 Introduction

The lack of fully comprehended pressure-state relationships based on classical biodiversity metrics led WGBIODIV to devise ToR1a: “Establish a sound theoretical basis relating variation in biodiversity metric values to changes in anthropogenic pressure on marine communities (e.g. incorporating components of community size and trophic structure into the derivation of biodiversity metrics, taking account of linkage to habitat types and consideration of spatial pattern).”

During the three-year working period from 2016 to 2018 several theoretical concepts for functional biodiversity indicators were introduced and explored (see Chapter 5.3). This chapter ‘Protocol on the development of theoretical concepts of biodiversity indicators (2016/2017)’ aims to distil generic principles for developing theory-based biodiversity indicators.

The focus of this protocol will be on the development of community indicators, as single-species indicators may require less theoretical understanding as they are in many cases linked more directly to human pressures. However, most parts of this protocol should be generically applicable to single-species indicators alike.

The steps are:

Step 1: Identify relevant features

Step 2: Develop theoretical underpinning

Step 3: Develop the indicator

Step 4: Establish pressure-state relationship(s) (for operational indicator)

Step 5: Obtain “adequate status” targets (for operational indicators)

Step 6: Perform indicator evaluation

Step 7: Apply indicator concept to field data

5.1.2 Background

Ecological indicators are essential in ecosystem assessments

The increasing awareness of broad scale impacts of human activities on the marine environment instigated the implementation of ecosystem based approaches to marine management, either within single management sectors (Link, 2010; Belgrano and Fowler, 2011; Hilborn, 2011; Link and Browman, 2014) or across the full range of managed human activities (Arkema *et al.*, 2006; Leslie and McLeod, 2007; ICES, 2015). The implementation of ecosystem based management approaches is commonly associated with either integrating multiple single elements (Ojaveer and Eero, 2011; Probst and Lynam, 2016) or full integrated ecosystem-level assessments (IEA); (Toth and Hizsnyik, 1998; Levin *et al.*, 2009; Möllmann *et al.*, 2014). In either case, integrated assessments are based on indicators (Garcia *et al.*, 2000; Jennings, 2005). Ecological indicators are intended to capture or represent relevant features of the ecosystem which representatively inform on wider aspects of ecosystem health thereby guiding management agencies towards targeted action (OECD, 1993; Helsenfeld and Enserink, 2008).

The implementation of ecosystem-based approaches to marine management has initiated intensive research and development programmes on environmental and ecological indicators (see for example Mace and Baillie, 2007; Helsenfeld and Enserink, 2008; EU-COM, 2010; Shin *et al.*, 2010; HELCOM, 2013). In fact, the number of suggested indicators has become overwhelmingly large, sometimes leading to redundancy in their content and meaning (Greenstreet *et al.*, 2012a; Shephard *et al.*, 2012). Thus, technical frameworks for selecting indicators based on quality criteria have been proposed and applied (Rice and Rochet, 2005; ICES, 2012; ICES, 2013; Probst *et al.*, 2014; Queirós *et al.*, 2016). These frameworks define evaluation criteria to indicators addressing (amongst others) the data quality, sensitivity and responsiveness towards changes in anthropogenic pressures, comprehensibility, theoretical underpinning and (cost) effectiveness.

In this chapter, we focus on two types of indicators: ‘operational indicators’ link human disturbances (pressure) to the state of an ecosystem component and ‘surveillance indicators’, which are used for surveillance of single ecosystem components without a clear assessment target and direct link to specified pressures (Shephard *et al.*, 2015). Surveillance indicators are not directly linked to specific pressures, but aim to warn management if the ecosystem is leaving previously known boundaries

The lack of theoretical underpinning leads to unclear status targets: The case of fish indicators

Classical fish biodiversity indicator metrics (species richness or evenness) fail to score well with regards to sensitivity and responsiveness towards human disturbances (Greenstreet, 2008). This is in large proportion due to the circumstance that the relationship

between human pressures and biodiversity indices, representing the ecological state, is often poorly understood (Rice, 2000). A study by Piet and Jennings (2005) revealed that several biodiversity indicators (e.g. Hill's diversity indices) did not show a consistent relationship with fishing intensity and concluded that a better theoretical understanding of the response of biodiversity metrics to anthropogenic pressures is necessary.

To address this problem, recent fish biodiversity indicators were based on the size-composition of communities, where marine communities' biotic interactions are considered to be driven by size-structure rather than by taxonomic distinction (Daan *et al.*, 2005; Greenstreet *et al.*, 2011). A prominent example for such a size-based biodiversity indicator is the Large Fish Indicator (LFI), which assesses the ratio between the biomass of small and large fish within a community (Greenstreet *et al.*, 2011; Shephard *et al.*, 2011). However, even for indicators such as the LFI, unexpected patterns in the pressure-state relationship emerged i.e. that the responsiveness (the time it takes for a state indicator to react to changes in pressure) of the LFI to fishing intensity was much longer than previously assumed (Fung *et al.*, 2013). This long-lagged responsiveness indicated that the initial theoretical concept of the LFI was complicated by trophic cascades within the fish community and that a deeper theoretical understanding of fishing impacts on fish communities may still be necessary. An equivalent type of species diversity indicator to the LFI does not currently exist for benthic communities. The majority of existing benthic diversity indicators are based on species abundance or biomass (e.g. the OSPAR common indicator "BH2 - Condition of benthic habitat defining communities (Multi-Metric Indices)") and thus knowledge of pressure-state relationships between benthic communities and anthropogenic impacts is increasingly necessary to support the development of these indices.

The lack of a clear, unambiguous understanding of theoretical concepts underlying ecological indicators can lead to difficulties in establishing assessment baselines for good environmental status (GES). Currently, many indicators are missing assessment targets that have been derived from theory. Instead assessment targets are usually based on historical evidence (Greenstreet *et al.*, 2011; Fock, 2014). In the lack of such historical evidence, time-series based approaches are suggested (Rochet *et al.*, 2010; Probst and Stelzenmüller, 2015). Both approaches, however, are associated with difficulties. Assessment targets established from historical evidence may become invalid in a changing environment e.g. if the targets become affected by climate change (ICES, 2015). Assessments based on time-series analyses do not provide meaningful baselines with regard to the true environmental status as they only inform on relative changes.

A well-established example for an assessment target based on a theoretical principle is the maximum sustainable yield (MSY) for exploited fish stocks (Jennings *et al.*, 2001). Originally the MSY-concept has been developed by Schaeffer based on the idea of surplus production (King, 2007). Surplus production describes an increased productivity of exploited populations when the population size is reduced (e.g. by fishing). The productivity of this population is enhanced because density-dependent inhibitions of growth are relieved. The MSY-concept has led to the development of reference points for fisheries management. Within the advisory framework of the International Council for the Exploration of the Sea (ICES) the MSY-principle is currently used to obtain limit values for sustainable rates of exploitation (F_{MSY}); (Lassen *et al.*, 2014). The MSY-principle has been

transferred and adapted to obtain reference points of sustainable impacts for endangered fish species (Le Quesne and Jennings, 2012) and benthic communities (Fock *et al.*, 2011).

5.1.3 A step-by-step approach to theory-based indicator conception

The following section describes several steps to develop a theoretical concept which can be used to design biodiversity indicators (Figure 5.1.1). This step-by-step guide focuses on state indicators which intend to capture aspects of biodiversity within communities of ecosystem components (e.g. fish, benthos or plankton). It remains to be explored if this approach is applicable to holistic ecosystem indicators and other types of indicators as well.

Step 1: Identify relevant features

Ecological indicators are eventually about environmental assessment and hence to be used in management context. Therefore, it is important to define the relevant features that the management framework is seeking to address. For example, the MSFD defines several ecological components and human pressures as relevant for the assessment of environmental status (see Annex III, Tables 1 & 2) and suggests potential indicators to be used for these assessments (see 2017/848/EU). However, most of the indicators are described only qualitatively and with varying degrees of detail. Hence, it has and still is taking huge efforts by scientists and political managers to come to terms on which exact ecological elements and human pressures to assess and how the indicators should be designed around these elements. Therefore, it is crucial to understand and agree upon which the relevant ecological features are reflected by the indicator metric and to which pressure they relate.

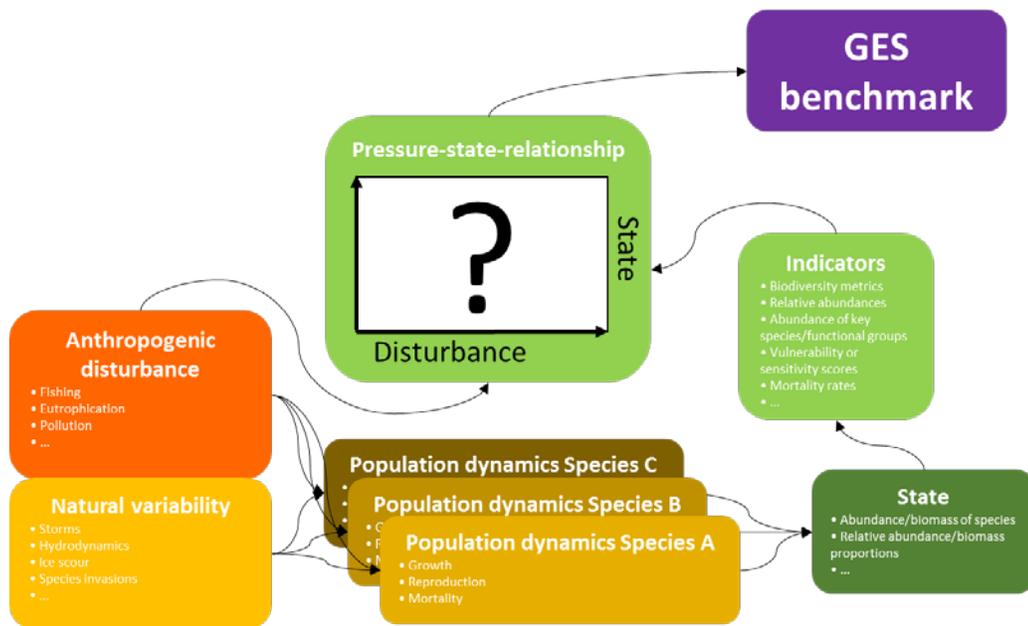


Figure 5.1.1. Linkages between anthropogenic disturbance, population dynamic processes, environmental states and biodiversity indicators. The ultimate goal is to derive assessment targets from the pressure-state relationship (PSR).

Step 2: Develop theoretical underpinning

The evaluation of OSPAR biodiversity indicators in 2013 by WGBIODIV revealed that many indicators scored poorly on the conceptual criterion to be “theoretically sound” (ICES, 2013). The lack of a theoretical underpinning of these indicators resulted in problems to define meaningful status targets and thus made the assessment of biodiversity elements based on these indicators difficult. Furthermore, classical indices of biodiversity provided ambiguous results with relation to human pressures and hence scored generally badly on evaluations by ICES and OSPAR (Greenstreet, 2008; Greenstreet *et al.*, 2011). WGBIODIV therefore concluded that it was essential to underlie biodiversity indicators with a sound theoretical concept that allows to formulate hypotheses on the pressure-state relationship *a priori* and allowing for rigorous testing of these hypotheses using empirical data.

In the following three types of theory-based indicators are described to demonstrate the way ecological theories can facilitate the design of biodiversity indicators.

Trait-based indicators

Species communities are characterised by the abundance of different species which influence their composition and diversity (Begon *et al.*, 1996). The abundance of each species within a community is influenced by abiotic conditions and biotic interactions, which in turn affect its population dynamics. Population dynamics are driven by processes, namely growth, reproduction and mortality (Jennings *et al.*, 2001), which in turn are depending on external environmental factors and intrinsic species-specific traits (Figure 5.1.1). The population dynamics of each species will determine the composition and structure of the community.

Knowing which traits render members of a community susceptible to specific pressures allows to build quantitative indices that capture and condense this sensitivity (McGill *et al.*, 2006; Gravel *et al.*, 2016). An example of such a trait-based indicator can be found in the concept of the WGBIODIV benthic response indicator (see chapter 5.3 of this report) or the sensitivity of demersal fish species (Greenstreet *et al.*, 2012b). In each case, biological traits (age-at-maturity, maximum length, burrowing depth, fragility, etc.) are combined into an index of sensitivity, which then can be calculated for samples of differing species compositions and related to influencing factors (Beauchard *et al.*, 2017).

Size-based indicators

Size-based indicators can be considered as a special form of trait-based indicators, as body size is considered to be the major trait affected by human pressure, e.g. by trawl fishing (HELCOM, 2017; OSPAR, 2017). A classic example is the OSPAR Large Fish Indicator (LFI), which reflects the proportion of large vs. small fish in the demersal fish community (Greenstreet *et al.*, 2011; Shephard *et al.*, 2011; Modica *et al.*, 2014). Other examples are the Typical Length (a size composition indicator, ICES, 2014) and Mean Maximum Length (a species composition indicator) of fish communities implemented within the recent OSPAR Intermediate Assessment (OSPAR, 2017). The theoretical underpinning in relation to pressure is that selective fishing alters the size-structure of fish communities by reducing the number of large species across the community as well as within single-species populations (Pauly *et al.*, 1998; Jennings *et al.*, 2002; Probst *et al.*, 2013b).

In relation to state, Jennings *et al.* (2007) found that body size was related to trophic level in fish in the North Sea at the community level (see also Reum *et al.*, 2015). Barnes *et al.* (2010) demonstrated a relationship between fish size and trophic transfer efficiency. Riede *et al.* (2011) demonstrated that log-mean body size was significantly related to trophic level in marine invertebrates, and ectotherm and endotherm vertebrates using data on multiple ecosystems. Model simulations by Rossberg *et al.* (2008) have demonstrated that in food webs where trophic interactions dominate over other interactions, large species at high trophic levels are highly sensitive to loss of diversity at lower trophic levels (ICES, 2014).

Indicators based on diversity metrics

Indicators can be based on plain diversity metrics such as the Margalef-Index (Borja *et al.*, 2009a). In the 2017 OSPAR Intermediate Assessment the Margalef-Index was used to assess the state of soft-bottom benthic habitats in the southern North Sea (OSPAR, 2017). The Margalef-Index is an abundance-weighted species richness metric and is supposed to be negatively related to several anthropogenic pressures such as pollution or organic enrichment. Diversity metrics are also used in combined multimetric indicators such as the M-AMBI, which uses Shannon-diversity and species richness in combination with a trait based approach (Borja *et al.*, 2009b). Dominance and diversity indices (Menhinick index and Hulburt index) are combined within the OSPAR PH3 "Changes in plankton diversity" indicator, which corresponds to a multimetric index focusing at structure, namely heterogeneity, diversity, and contributions of each taxa to community diversity. By applying the Local contributions to beta diversity (LCBD, Legendre & De Caceres 2013) approach which uses variance in taxa distribution among sampling units, PH3 allows the identification of atypical community structures which can be considered for index calibration for future assessments or, instead, correspond to degraded areas in need of restoration (Budria *et al.*, 2017).

In some cases diversity metrics have been proven themselves as difficult to assess the impact of pressures, e.g. fishing (Piet and Jennings, 2005). Indicators of species richness or evenness depend very much on the sampling strategy i.e. the number of samples necessary to capture the true values of such metrics (Greenstreet and Piet, 2008). Hence, these metrics can be data-intensive, methodologically demanding and difficult to calculate and wrong applications can make their interpretation difficult (Greenstreet, 2008).

Alternatively, indicators can be based on knowledge, e.g. evidence from scientific literature, direct observations or logical, yet descriptive conclusions. However, we would like to distinguish this type of underpinning from the conceptual and theoretical underpinning as described above, as WGBIODIV concluded that this type of indicator rationale eventually will lead to undefinable assessment targets (ICES, 2013).

Step 3: Develop the indicator

At this step, it becomes necessary to decide on concrete indicator metric(s) to be calculated from available data. Effort should be spent on assessing which metric may be best suited to capture the concept of the indicator e.g. by analysing which metric would be most sensitive and specific to the relevant pressure(s) (see Greenstreet *et al.*, 2011; Probst and Oesterwind, 2014 for examples). Effort may also be needed to develop a meaningful pressure indicator metric (see Greenstreet *et al.*, 2011 for an example on communal fish-

ing pressure). Furthermore, the exact ecosystem components, which should be included into the metric calculation need to be defined. For example, some species may not be monitored well enough by a specific monitoring program (e.g. wide-ranging species like basking shark in demersal fisheries surveys) or may not be sensitive to the impact of certain human pressures (e.g. pelagic fish to demersal trawling). To make the indicator metric as suitable to the concept as possible, careful decisions have to be made regarding all relevant aspects, e.g. the suite of included features and traits, the type of metric to calculate or which cut-off threshold to choose.

The underlying data that will be used to calculate the indicator most likely will have to be compiled, validated and quality assured. This can be a lengthy and time consuming process (Moriarty *et al.*, 2017). In fact, the completeness and quality of regional data bases is diverse and in many cases it has been proven as challenging to gather the necessary at the relevant scales (e.g. compile vessel-monitoring data or observers-at-sea data across all EU member states). Hence, this step is very much about cleaning and consolidating the existing data, correcting or eradicating erroneous entries as well as filtering the data to include only the relevant spatial units and ecological elements.

Step 4: Establish a pressure–state relationship (PSR)

This step is necessary, if the intended indicator is supposed to become a fully operational assessment indicator. Contrary, surveillance indicators do not need to be linked explicitly to pressure(s) (Shephard *et al.*, 2015) and step 4 may be disregarded.

In the ideal case, mathematical equation(s) define the relationship between pressure and states. These pressure-state relationships (PSR) can thus be described as theoretically formalised, conceptually validated or conceptual non-validated (Table 5.1.1). An example for a theoretically formalised PSR is the relationship between fishing mortality and yield-per-recruit (Beverton and Holt, 1957; Jennings *et al.*, 2001). A conceptual validated PSR is tested using empirical data (Fock *et al.*, 2011; Large *et al.*, 2013) or can use modelling to validate and obtain the PSR (Fulton *et al.*, 2005; Le Quesne and Jennings, 2012; Probst *et al.*, 2013b).

If possible, the PSR should not only indicate the direction of the impact, but also the magnitude and form of the relationship (linear, asymptotic, hump-shaped, bimodal, hockey-stick, etc.) (Samhuri *et al.*, 2010). The knowledge on the form of the PSR is important for step 5.

Table 5.1.1 Types of pressure-state relationships (PSR)

PSR-type	Description	Examples
Theoretically formalized	The relationship between pressure(s) and state can be analytically derived from equations	Fishing mortality vs. yield per recruit
Conceptually validated	Conceptual PSR is postulated based on and validated by empirical data	WGBIODIV benthic response indicator
Conceptually non-validated	Conceptual PSR is postulated, but not validated	Trawling frequency vs. benthic disturbance indicator

Step 5: Obtain “adequate status” (assessment) targets

If the indicator concept is based on a theoretical framework which connects pressure and states in fully quantitative equations, targets for GES should be obtainable from these equations, if it possible to solve for local minima, maxima or inflection points (see below and Figure 5.1.2). Otherwise operating models or empirical data can help to determine GES thresholds by providing pressure-state relationships (PSR).

Due to their “alarm” function, surveillance indicators do not need strict theoretical underpinning. The designation of status targets is therefore less difficult for this type of indicator and can be obtained from values observed in the time-series (Probst and Stelzenmüller, 2015; Shephard *et al.*, 2015).

Depending on the form of the PSR it may be possible to determine benchmarks for GES (Samhuri *et al.*, 2010). If the PSR is non-linear, inflection or breakpoints may be used to identify where a minimal change in pressure will lead to a disproportional change in state (Figure 5.1.2). However, it is notable that this is only valid for certain types of PSR that contain some sort of turning or break point. In other cases, the GES-benchmark may be found by taking other relevant ecological features into account. An example from classical fisheries management: Fishing mortality (F) has a monotonous logarithmic relationship to the cohort number (N) (similar to Figure 5.1.2C). Hence, no point is discernible at which a small increase in F has disproportionately high impact on N. But as the cohort number relates to spawning stock biomass, which in turn is related to recruitment (and thus productivity) of the stock, thresholds for SSB (and indirectly N and F) can be derived from the spawner-recruit relationship.

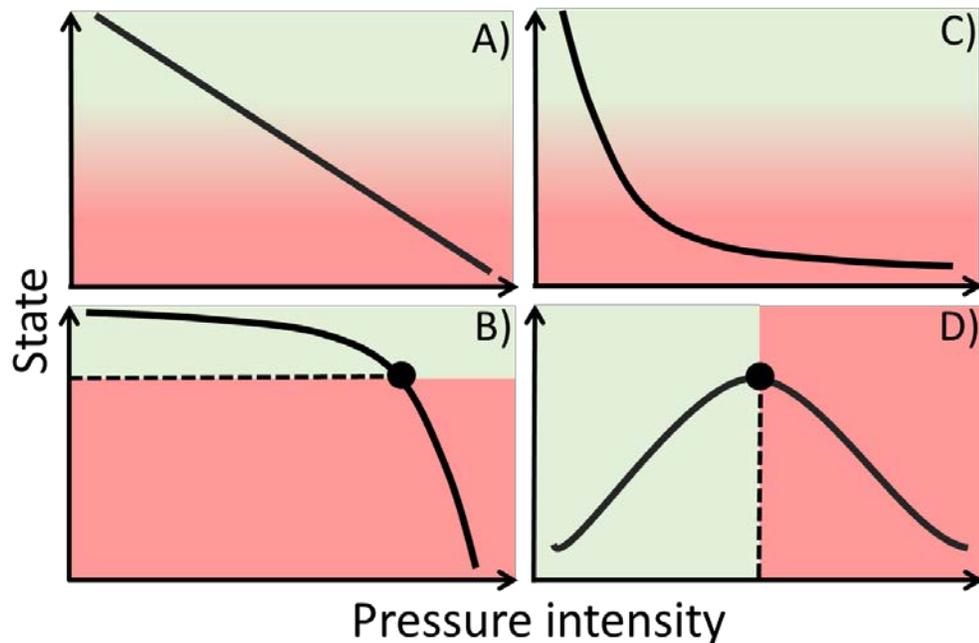


Figure 5.1.2. Types of pressure-state relationships (PSR) and their potential for determining assessments thresholds. A linear PSR does not allow the mathematical determination of an assessment threshold (A), an asymptotic relationship allows the determination of an inflection point, but an inflection point may only be a valid benchmark for good environmental status (GES) in a resistant system (B), whereas in systems with low resistance any amount of pressure will lead to drastic declines in environmental state (C). A hump-shaped relationship could inform sustainable levels of pressure (D). Colours indicate assessment outcome, green='good', red='bad', black dots indicate inflection points or maxima.

At best, status or assessment targets should account for uncertainty e.g. by adding precautionary buffers to the determined reference points. Again, the relationship between spawners and recruits of exploited fish stocks can provide an example: The turning-point in the hockey-stick relationship between spawners and recruits is defined as minimum sustainable biomass (B_{lim}). Adding a safety buffer of 40% to B_{lim} yields the precautionary reference point (B_{pa}) (ICES, 2017). Accordingly, any other quantitative measures of uncertainty can be used i.e. a given quantile of the stochastic distribution of an indicator metric (Probst *et al.*, 2013a; Probst, 2017).

Step 6: Perform indicator evaluation

In this final step the quality of the indicator is assessed to decide whether the indicator is usable within environmental assessment frameworks such as the MSFD or regional seas conventions (Rice and Rochet, 2005; ICES, 2015). Building on work by WGECO (ICES, 2012), WGBIODIV introduced a scoring procedure for quantifying expert judgement (ICES, 2013), which can be generically applied to biological indicators (Probst *et al.*, 2014). The WGBIODIV criteria are based on criteria defined by Rice and Rochet (2005) and Kershner *et al.* (2011) and cover criteria such as data availability, spatial coverage, pressure-specificity and responsiveness, comprehensibility, cost-effectiveness and theoretical soundness (ICES, 2013). Independent of the framework that will be used, indicator eval-

uation is helpful to identify the strengths and gaps of indicators promoting their improvement and increasing their acceptance within the scientific and policy community.

Step 7: Do the assessment

Once the indicator concept, the PSR and the assessment targets have been established, it is time to take the indicator to the real world. Similar to step 4 and step 5, this step only applies to operational indicators and is about dealing with unexpected problems when transferring hypothetical concepts into scientific applications intended to inform management and policy.

After calculating the indicator, establishing the PSR and assessment target, the actual state of the indicator can be assessed. What sounds like a straightforward task, however, can turn into a lengthy process in which the assessment result has to be validated, interpreted and communicated.

The OSPAR Intermediate Assessment 2017 is a good example of how existing and new indicators have been adapted to fit to the actual assessment needs. For example, several fish indicators such as the 'Typical Length' and the 'Mean Maximum Length' have been calculated for sub-divisions as the fish communities have been found to be heterogeneous within the OSPAR regions (e.g. the Celtic Seas) and reporting by OSPAR regions would have masked sub-divisional trends. The Large Fish Indicator for the North Sea was originally based on the IBTS-NS Quarter-1 survey, but has been calculated based on several surveys of the North Sea to confirm the observed trajectory of the original indicator time-series.

Another important aspect is how the result of the assessment is communicated to managers and decision makers. Thus, it is important to consider the way assessment results are disseminated. It is certainly important that the indicator metric as well as its assessment targets are fairly easy to understand and to interpret (Rochet and Rice, 2005; ICES, 2013). ICES adopted a traffic light approach to indicate the status of fish stock indicators. Contrary, the current OSPAR 2017 Intermediate assessment does not have a unified approach to present the results and the current versions of the indicator fact sheets do not present a straightforward assessment result.

After performing steps 1–6, operational indicators in most cases will run through iterative adaptation processes until the stakeholder needs for spatial and temporal resolution, understanding, standardisation and acceptance are met.

5.1.4 Case studies

This section analyses how the seven-step-framework applies to the WGBIODIV benthic response indicator as well as to three indicators by OSPAR and HELCOM. The aim of this section is to analyse how well these indicators fit to the WGBIODIV development protocol and what caveats and gaps may result from missing important steps of this protocol.

Table 5.1.2. Overview on the congruency of the development procedure of several ecosystem indicators against the WGBIODIV development protocol.

Name	OSPAR Large Fish Indicator LFI	WGBIODIV Benthic response indicator	HELCOM abundance of coastal fish functional groups	OSPAR Abundance and distribution of cetaceans
Indicator type	Operational indicator	Operational indicator	Operational indicator	Surveillance indicator
<i>Step 1. Identify relevant features</i>	Pressure(s): Community fishing pressure State: Proportion of large fish in demersal fish community	Pressure(s): Demersal trawling State: Composition of benthic community	Pressure(s): Eutrophication, fishing, habitat degradation, climate change State: Abundance of trophic fish guilds	Pressure(s): None State: Abundance and distribution of several cetacean species
<i>Step 2. Develop theoretical underpinning</i>	Conceptual	Conceptual	Conceptual	Knowledge-based
<i>Step 3. Develop the indicator</i>	Calculate the ratio between biomass of small and large fish Define threshold to separate small from large fish by statistical procedure Filter Survey data for demersal species and standard survey area	Define biological traits which are assumed sensitive to fishing Develop indicator metric combining these traits;	Calculate abundance estimates of piscivores and cyprinids/mesopredators during the period 2011–2015 Develop reference conditions based on previous values in the time-series	Calculate abundance estimates Create distribution maps Compare changes in abundance
<i>Step 4. Establish a pressure-state relationship</i>	Conceptual validated	Conceptual validated	Conceptual non-validated	None

<i>Step 5. Obtain "adequate status" (assessment) targets for operational indicators</i>	Historical	Pending	Time-series based	Time-series based
<i>Step 6. Perform indicator evaluation</i>	Done, evaluated as operational	Pending	Done, evaluated as operational	Pending
<i>Step 7. Apply indicator concept to field data</i>	Greater North Sea; Celtic Seas; Bay of Biscay	Case study in the Dutch EEZ and Bay of Biscay	Across multiple coastal HELCOM assessment units in the Baltic Sea	Depending on species data was available for the North Sea, Celtic Sea, Bay of Biscay and Iberian Coast
Potential improvements	Define length threshold for large fish based on food web theory	Establish assessment target based on PSR	Validate PSR e.g. eutrophication vs. ratio of cyprinids/mesopredators	Use population viability analysis to obtain estimates for extinction risk and minimum viable population sizes

The OSPAR LFI

The OSPAR Large Fish Indicator (LFI) is currently one of the few ecosystem indicators that has been developed, validated and improved and thus is considered as operational (Greenstreet *et al.*, 2011; ICES, 2015; OSPAR, 2017). The LFI is based on two theoretical concepts, namely size-structured food webs and size-selective fishing. However, these concepts were not sufficiently mathematically formalised to close some gaps that became apparent throughout the development of the LFI.

Firstly, the definition of a size threshold to distinguish small from large fish was not derived from food web theory and impacts of size-selective fishing. Instead, all applied case studies of the LFI in the North Sea, the Celtic Sea and Bay of Biscay (Greenstreet *et al.*, 2011; Shephard *et al.*, 2011; Modica *et al.*, 2014) applied a technical procedure fitting polynomial smoothers to the LFI time-series choosing the length threshold, which produced the best smoother fit.

Secondly, the theoretical foundation of the LFI was not sufficient to designate a threshold for good environmental status. This threshold had to be determined based on historical time-series from fisheries surveys dating back to the 1920s (Greenstreet *et al.*, 2011). The historical approach to set a GES-threshold for the LFI was adapted for the Celtic Sea (Shephard *et al.*, 2011) and a statistical procedure was used for the LFI of the Bay of Biscay (Modica *et al.*, 2014). Thus, the current theoretical foundation of the LFI is not suffi-

cient to go beyond the operationalisation procedures suggested by Probst and Stelzenmüller (2015).

Finally, the temporal dynamics of the pressure-state relationships indicate that the underlying theoretical concepts of the LFI are more complicated than originally assumed. The LFI responds with time lags of eight to sixteen years to changes in fishing pressure, and simulation studies suggest that complex food web interactions result in trophic cascades, which may need decades until equilibrium conditions are reached (Fung *et al.*, 2013; Shephard *et al.*, 2013)

In conclusion, the LFI is well rooted in theoretical concepts and its validity is confirmed by multiple studies in different marine regions. However, the theoretical foundation of the LFI is qualitative and not quantitative and hence the establishment of GES-thresholds has been based on historical records.

OSPAR abundance and distribution of cetaceans

The indicator “abundance and distribution of cetaceans” is a common OSPAR indicator applied to all OSPAR regions and contributing to both descriptors D1 and D4 (OSPAR, 2017). It is based on the rationale that cetaceans as top predators form an important part of marine biodiversity and are affected by multiple human pressures (incidental bycatch, collisions with ships, underwater noise by shipping or seismic activities, prey depletion by overfishing, habitat loss/degradation, pollution, marine debris, climate change). Thus, high cetacean abundance and distribution is assumed to be related to a good environmental status.

Currently no GES-thresholds are defined, mainly due to a lack of sufficient monitoring data, which make precise estimates of abundance and distribution trends difficult. Furthermore, the PSR are not very specific and given the weak data availability empirical support for PSR cannot be provided. For the same reason, formal pressure-state relationships cannot be provided. While high abundances of cetaceans may indicate a good environmental status, their absence does thus not allow determining the actual status of an ecosystem and driving anthropogenic causes due to the low data availability.

Thus, the assessment is rather qualitative (based on expert judgement) than quantitative. In conclusion, an evaluation of this indicator is not possible at the moment as it requires better data support and further methodological development.

This indicator could gain from applying population viability analysis to single populations in order to estimate minimum viable population sizes and extinction risks under prevailing conditions (Boyce, 1992; Traill *et al.*, 2007).

HELCOM abundance of coastal fish functional groups

This HELCOM core indicator is part of the State of the Baltic Sea Holistic Assessment 2017 (HELCOM, 2017) and contributes to the assessment of food webs under criterion D4C2. The indicator evaluates the abundance of two trophic fish guilds, piscivores and cyprinids (or, if not applicable, mesopredators), in coastal regions of the Baltic Sea. The rationale for this indicator is that human pressures such as eutrophication, fishing and habitat degradation reduce the abundance of top predators (piscivores) while improving the conditions for cyprinids/mesopredators (Persson *et al.*, 1991; Sandström and Karas, 2002; Bergström *et al.*, 2016). Thus, obtaining a good environmental status requires that

both abundance of piscivores is above a particular threshold and abundance of cyprinids/mesopredators is within a particular range.

Site-specific thresholds/ranges are determined in reference to baseline conditions derived from long-term time-series (≥ 15 years). However, these reference conditions include uncertainty regarding the real status of the fish guilds, as this was not derived from ecological considerations but from the existent time-series only. In addition, the thresholds/ranges are not available for all areas due to a lack of data i.e. shortness of the available time-series. Finally, environmental gradients in the Baltic Sea are steep and the existing monitoring programmes are concentrated in the Northern and Eastern parts of the Baltic Sea. The baselines obtained from these areas may not be applicable to other coastal areas of the Baltic Sea where coastal fish monitoring programmes have not yet established.

5.1.5 Conclusions

This non-exhaustive review of four biodiversity indicators illustrates that the establishment of fully theoretically formalised indicators is very challenging. At the moment, the most advanced theoretical underpinning can be found in indicators, which have a long tradition of assessing the sustainability of exploitation of fish stocks. Compared to these indicators (fishing mortality F and spawning stock biomass SSB), the indicators established by the regional seas conventions are rather new. Maybe due to their relatively young age, none of the RSC biodiversity indicators is rooted in a fully formalized theoretical framework. While some indicators are rooted in ecological concepts, even in the best cases their conceptualisation is not fully numerical and does not allow to derive theory-based pressure-state relationships and assessment targets. Instead, empirical data are used to validate and define pressure state relationships and assessment targets are derived from historical data or by statistical methods. Many of the species-specific indicators (and their higher-levels composite products) could overcome these shortcomings by using population modelling based on species-specific traits. Improved monitoring programs would help to calibrate and validate the outcomes of such population models.

5.1.6 References

- Arkema, K. K., Abramson, S. C., and Dewsbury, B. M. 2006. Marine ecosystem-based management: from characterization to implementation. *Frontiers in Ecology and the Environment*, 4: 525–532.
- Barnes, C., Maxwell, D. L., Reumann, D. C., and Jennings, S. 2010. Global patterns in predator-prey size relationships reveal size dependency of trophic transfer efficiency. *Ecology*, 91: 222–232.
- Beauchard, O., Veríssimo, H., Queirós, A. M., and Herman, P. M. J. 2017. The use of multiple biological traits in marine community ecology and its potential in ecological indicator development. *Ecological Indicators*, 76: 81–96.
- Begon, M. E., Harper, J. L., and Townsend, C. R. 1996. *Ecology*, Blackwell Science, Oxford.
- Belgrano, A., and Fowler, C. W. 2011. *Ecosystem-based management for marine fisheries - An evolving perspective*, Cambridge University Press, New York. 385 pp.
- Bergström, L., Heikinheimo, O., Svirgsden, R., Kruze, E., Ložys, L., Lappalainen, A., Saks, L., *et al.* 2016. Long term changes in the status of coastal fish in the Baltic Sea. *Estuarine, Coastal and Shelf Science*, 169: 74–84.

- Beverton, R. J. H., and Holt, S. 1957. On the dynamics of exploited fish populations, Ministry of Agriculture, Fisheries and Food, London.
- Borja, A., Miles, A., Occhipinti-Ambrogi, A., and Berg, T. 2009a. Current status of macroinvertebrate methods used for assessing the quality of European marine waters: implementing the Water Framework Directive. *Hydrobiologia*, 633: 181–196.
- Borja, A., Muxika, I., and Rodríguez, J. G. 2009b. Paradigmatic responses of marine benthic communities to different anthropogenic pressures, using M-AMBI, within the European Water Framework Directive. *Marine Ecology*, 30: 214–227.
- Boyce, M. S. 1992. Population viability analysis. *Annual Review of Ecology and Systematics*, 23: 481–506.
- Budria, A., Aubert, A., Rombouts, I., Ostle, C., Atkinson, A., Widdicombe, C., Goberville, E., Artigas, F., Johns, D., Padegimas, B., Corcoran, E. and McQuatters-Gollop, A. (2017). Cross-linking plankton indicators to better define GES of pelagichabitats EcAprHA Deliverable WP1.4. OSPAR, London, UK, 54.
- Daan, N., Gislason, H., Pope, J., and Crice, J. 2005. Changes in the North Sea fish community: evidence of indirect effects of fishing? *ICES Journal of Marine Science*, 62: 177–188.
- EU-COM 2010. Commission decision of 1 September 2010 on criteria and methodological standards on good environmental status of marine waters (notified under document C(2010) 5956) (2010/477/EU). In 2010/477/EU, pp. 12–24. Ed. by E. Commission. European Commission.
- Fock, H. O. 2014. Estimating historical trawling effort in German Bight from 1924 to 1938. *Fisheries Research*, 154: 26–37.
- Fock, H. O., Kloppmann, M., and Stelzenmüller, V. 2011. Linking marine fisheries to environmental objectives: a case study on seafloor integrity under European maritime policies. *Environmental Science & Policy*, 14: 289–300.
- Fulton, E. A., Smith, A. D. M., and Punt, A. E. 2005. Which ecological indicators can robustly detect effects of fishing? *ICES Journal of Marine Science*, 62: 540–551.
- Fung, T., Farnsworth, K. D., Shephard, S., Reid, D. G., and Rossberg, A. G. 2013. Why the size structure of marine communities can require decades to recover from fishing. *Marine Ecology Progress Series*, 484: 155–171.
- Garcia, S. M., Staples, D. J., and Chesson, J. 2000. The FAO guidelines for the development and use of indicators for sustainable development of marine capture fisheries and an Australian example of their application. *Ocean & Coastal Management*, 43: 537–556.
- Gravel, D., Albouy, C., and Thuiller, W. 2016. The meaning of functional trait composition of food webs for ecosystem functioning. *Philos Trans R Soc Lond B Biol Sci*, 371.
- Greenstreet, S. P. R. 2008. Biodiversity of North Sea fish: why do the politicians care but marine scientists appear oblivious to this issue? *ICES Journal of Marine Science*, 65: 1515–1519.
- Greenstreet, S. P. R., Fraser, H. M., Rogers, S. I., Trenkel, V. M., Simpson, S. D., and Pinnegar, J. K. 2012a. Redundancy in metrics describing the composition, structure, and functioning of the North Sea demersal fish community. *ICES Journal of Marine Science*, 69: 8–22.
- Greenstreet, S. P. R., and Piet, G. J. 2008. Assessing the sampling effort required to estimate α species diversity in the groundfish assemblages of the North Sea. *Marine Ecology Progress Series*, 364: 181–197.
- Greenstreet, S. P. R., Rogers, S. I., Rice, J. C., Piet, G. J., Guirey, E. J., Fraser, H. M., and Fryer, R. J. 2011. Development of the EcoQO for the North Sea fish community. *ICES Journal of Marine Science*, 68: 1–11.

- Greenstreet, S. P. R., Rossberg, A. G., Fox, C., Le Quesne, W. J. F., Blasdale, P., Boulcott, P., Mitchell, I., *et al.* 2012b. Demersal fish biodiversity: species-level indicators and trends-based targets for the marine strategy framework directive. *ICES Journal of Marine Science*, 69: 1789–1801.
- HELCOM. 2013. HELCOM core indicators - Final report of the HELCOM CORSET project. 136. 74 pp.
- HELCOM. 2017. First version of the 'State of the Baltic sea' report - June 2017 - to be updated in 2018.
- Helsenfeld, P., and Enserink, E. L. 2008. OSPAR ecological quality objectives: the utility of health indicators in the North Sea. *ICES Journal of Marine Science*, 65: 1392–1397.
- Hilborn, R. 2011. Future directions in ecosystem based fisheries management: A personal perspective. *Fisheries Research*, 108: 235–239.
- ICES. 2012. Report of the working group on ecosystem effects of fishing activities (WGECO) 26. 192 pp.
- ICES. 2013. Report of the working group of biodiversity science (WGBIODIV). 02. 64 pp.
- ICES. 2014. Report of the working group on the ecosystem effects of fishing activities (WGECO). 26. 174 pp.
- ICES. 2015. Report of the working group on biodiversity science (WGBIODIV). 04. 308 pp.
- ICES 2017. ICES Advice 2017. *In* Book 6. Ed. by ICES. ICES, Copenhagen.
- Jennings, S. 2005. Indicators to support an ecosystem approach to fisheries. *Fish & Fisheries*, 6: 212–232.
- Jennings, S., De Oliveira, J. A. A., and Warr, K. J. 2007. Measurement of body size and abundances in tests of macroecological and food web theory. *Journal of Animal Ecology*, 76: 72–82.
- Jennings, S., Greenstreet, S. P. R., Hill, L., Piet, G. J., Pinnegar, J. K., and Warr, K. J. 2002. Long-term trends in the trophic structure of the north sea fish community: Evidence from stable-isotope analysis, size-spectra and community metrics. *Marine Biology*, 141: 1085–1097.
- Jennings, S., Kaiser, M. J., and Reynolds, J. D. 2001. *Marine Fisheries Ecology*, Blackwell Science, Oxford. 432 pp.
- Kershner, J., Samhouri, J. F., James, C. A., and Levin, P. S. 2011. Selecting indicator portfolios for marine species and food webs: A Pudget Sound case study *PLoS ONE*, 6: e25248.
- King, M. 2007. *Fisheries Biology, Assessment and Management*. p. 382 pp. Blackwell Publishing, Oxford.
- Large, S. I., Fay, G., Friedland, K. D., and Link, J. S. 2013. Defining trends and thresholds in responses of ecological indicators to fishing and environmental pressures. *ICES Journal of Marine Science*, 70: 755–767.
- Lassen, H., Kelly, C., and Sissenwine, M. 2014. ICES advisory framework 1977–2012: from Fmax to precautionary approach and beyond. *ICES Journal of Marine Science*, 71: 166–172.
- Le Quesne, W. J. F., and Jennings, S. 2012. Predicting species vulnerability with minimal data to support rapid risk assessment of fishing impacts on biodiversity. *Journal of Applied Ecology*, 49: 20–28.
- Legendre P., De Cáceres M. 2013. Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecology Letters* 16: 951–963
- Leslie, H. M., and McLeod, K. L. 2007. Confronting the challenges of implementing marine ecosystem-based management. *Frontiers in Ecology and the Environment*, 5: 540–548.

- Levin, P. S., Fogarty, M. J., Murawski, S. A., and Fluharty, D. 2009. Integrated ecosystem assessments: developing the scientific basis for ecosystem-based management of the ocean. *PLoS Biology*, 7: 23–28.
- Link, J. 2010. *Ecosystem-based fisheries management: Confronting tradeoffs*, Cambridge University Press. 224 pp.
- Link, J. S., and Browman, H. I. 2014. Integrating what? Levels of marine ecosystem-based assessment and management. *ICES Journal of Marine Science*, 71: 1170–1173.
- Mace, G. M., and Baillie, J. E. 2007. The 2010 biodiversity indicators: challenges for science and policy. *Conservation Biology*, 21: 1406–1413.
- McGill, B. J., Enquist, B. J., Weiher, E., and Westoby, M. 2006. Rebuilding community ecology from functional traits. *Trends Ecol Evol*, 21: 178–185.
- Modica, L., Velasco, F., Preciado, I., Soto, M., and Greenstreet, S. P. R. 2014. Development of the large fish indicator and associated target for a Northeast Atlantic fish community. *ICES Journal of Marine Science*, 71: 2403–2415.
- Möllmann, C., Lindegren, M., Blenckner, T., Bergstrom, L., Casini, M., Diekmann, R., Flinkman, J., et al. 2014. Implementing ecosystem-based fisheries management: from single-species to integrated ecosystem assessment and advice for Baltic Sea fish stocks. *ICES Journal of Marine Science*, 71: 1187–1197.
- Moriarty, M., Greenstreet, S. P. R., and Rasmussen, J. 2017. Derivation of groundfish survey monitoring and assessment data products for the Northeast Atlantic area. 8. 240 pp.
- OECD. 1993. OECD core set of indicators for environmental performance reviews. ICES Document OCDE/GD(93)179: 83. 39 pp.
- Ojaveer, H., and Eero, M. 2011. Methodological challenges in assessing the environmental status of a marine ecosystem: case study of the Baltic Sea. *PLoS ONE*, 6: e19231.
- OSPAR. 2017. Intermediate Assessment 2017.
- Pauly, D., Christensen, V., Dalsgaard, J., Froese, R., and Torres, F. 1998. Fishing down marine food webs. *Science*, 279: 860–863.
- Persson, L., Diehl, S., Johansson, L., Andersson, G., and Hamrin, S. F. 1991. Shifts in fish communities along the productivity gradient of temperate lakes: patterns and the importance of size-structured interactions. *Journal of Fish Biology*, 38: 281–294.
- Piet, G., and Jennings, S. 2005. Response of potential fish community indicators to fishing. *ICES Journal of Marine Science*, 62: 214–225.
- Probst, W. N. 2017. A generic aggregation approach to account for statistical uncertainty when combining multiple assessment results. *Ecological Indicators*, 73: 686–693.
- Probst, W. N., Kloppmann, M., and Kraus, G. 2013a. Indicator-based assessment of commercial fish species in the North Sea according to the EU Marine Strategy Framework Directive (MSFD). *ICES Journal of Marine Science*, 70: 694–706.
- Probst, W. N., and Lynam, C. P. 2016. Aggregated assessment results depend on aggregation method and framework structure - a case study within the European Marine Strategy Framework Directive. *Ecological Indicators*, 61: 871–881.
- Probst, W. N., and Oesterwind, D. 2014. How good are alternative indicators for spawning-stock biomass (SSB) and fishing mortality (F)? *ICES Journal of Marine Science*, 71: 1137–1141.

- Probst, W. N., Rau, A., Diekmann, R., Von Dorrien, C., Seidel, H., Fock, H., Kraus, G., *et al.* 2014. Eine Thünen-Evaluierung von fisch- und fischereibezogenen Indikatoren der EU Meeresstrategie-Rahmenrichtlinie (MSRL). 25.
- Probst, W. N., and Stelzenmüller, V. 2015. A benchmarking and assessment framework to operationalise ecological indicator based on time-series analysis. *Ecological Indicators*, 55: 94–106.
- Probst, W. N., Stelzenmüller, V., and Kraus, G. 2013b. A simulation-approach to assess the size structure of commercially exploited fish populations within the European Marine Strategy Framework Directive. *Ecological Indicators*, 24: 621–632.
- Queirós, A. M., Strong, J. A., Mazik, K., Carstensen, J., J., B., Somerfield, P. J., Bruhn, A., *et al.* 2016. An objective framework to test the quality of candidate indicators of good environmental status. *Frontiers in Marine Science*.
- Reum, J. C., Jennings, S., and Hunsicker, M. E. 2015. Implications of scaled delta15N fractionation for community predator-prey body mass ratio estimates in size-structured food webs. *J Anim Ecol*, 84: 1618–1627.
- Rice, J. 2000. Evaluating fishery impacts using metrics of community structure. *ICES Journal of Marine Science*, 57: 682–688.
- Rice, J., and Rochet, M. 2005. A framework for selecting a suite of indicators for fisheries management. *ICES Journal of Marine Science*, 62: 516–527.
- Riede, J. O., Brose, U., Ebenman, B., Jacob, U., Thompson, R., Townsend, C. R., and Jonsson, T. 2011. Stepping in Elton's footprints: a general scaling model for body masses and trophic levels across ecosystems. *Ecology Letters*, 14: 169–178.
- Rochet, M.-J., and Rice, J. C. 2005. Do explicit criteria help in selecting indicators for ecosystem-based fisheries management? *ICES Journal of Marine Science*, 62: 528–539.
- Rochet, M.-J., Trenkel, V. M., Carpentier, A., Coppin, F., De Sola, L. G., Léauté, J.-P., Mahé, J.-C., *et al.* 2010. Do changes in environmental and fishing pressures impact marine communities? An empirical assessment. *Journal of Applied Ecology*, 47: 741–750.
- Rossberg, A. G., Ishii, R., Amemiya, T., and Itoh, K. 2008. The top-down mechanism for body-mass-abundance scaling. *Ecology*, 89: 567–580.
- Samhuri, J. F., Levin, P., S., and Ainsworth, C. H. 2010. Identifying thresholds for ecosystem-based management. *PLoS ONE*, 5: e8907.
- Sandström, A., and Karas, P. 2002. Effects of eutrophication on young-of-the-year freshwater fish communities in coastal areas of the Baltic. *Environmental Biology of Fishes*, 63: 89–101.
- Shephard, S., Fung, T., Houle, J. E., Farnsworth, K. D., Reid, D. G., and Rossberg, A. G. 2012. Size-selective fishing drives species composition in the Celtic Sea. *ICES Journal of Marine Science*, 69: 223–234.
- Shephard, S., Fung, T., Rossberg, A. G., Farnsworth, K. D., Reid, D. G., Greenstreet, S. P. R., and Warnes, S. 2013. Modelling recovery of Celtic Sea demersal fish community size-structure. *Fisheries Research*, 140: 91–95.
- Shephard, S., Greenstreet, S. P. R., Piet, G. J., Rindorf, A., and Dickey-Collas, M. 2015. Surveillance indicators and their use in implementation of the Marine Strategy Framework Directive. *ICES Journal of Marine Science*, 72: 2269–2277.
- Shephard, S., Reid, D. G., and Greenstreet, S. P. R. 2011. Interpreting the large fish indicator for the Celtic Sea. *ICES Journal of Marine Science*, 68: 1963–1972.

- Shin, Y.-J., Shannon, L. J., Bundy, A., Coll, M., Aydin, K., Bez, N., Blanchard, J. L., *et al.* 2010. Using indicators for evaluating, comparing, and communicating the ecological status of exploited marine ecosystems. 2. Setting the scene. *ICES Journal of Marine Science*, 67: 692–716.
- Toth, F. L., and Hizsnyik, E. 1998. Integrated environmental assessment methods: Evolution and applications. *Environmental Modeling and Assessment*, 3: 192–207.
- Traill, L., Bradshaw, C., and Brook, B. 2007. Minimum viable population size: A meta-analysis of 30 years of published estimates. *Biological Conservation*, 139: 159–166.

5.2 Standardizing sampling design and analysis to obtain unbiased estimates of species diversity (ToR1b, Deliverable 2)

Deliverable 2: Combined analysis and review on impacts of sampling size on performance of biodiversity metrics (2016–2018)

5.2.1 Introduction

The quantification of changes in biodiversity is often considered central to studies of human impacts on marine ecosystems, but metrics of biodiversity have proven exceptionally difficult to quantify reliably. For the sake of simplicity, the following chapter will be limited to the metric of richness, whether taxonomic or functional. The difficulties to reliably quantify richness are not only related to problems involved in species identification or functional classification. Many of the problems arise because the statistical properties of richness estimators are poorly known and because the richness estimates are based on diverging methods and sampling procedures.

Richness estimates are curtailed by the number of individuals sampled, because if only x individuals are caught, only x species can be identified. Drawing samples from a diverse assemblage, richness will increase with the number of individuals sampled and with the number of samples collected. Ideally, the relationship is asymptotic and a value of total richness characterizing the assemblage in a delimited area at a given time can be derived (Figure 5.2.1). But species accumulation curve can deviate from the ideal shape, if species migrate across the borders of the investigated area or when samples have been collected over a period of time.

Because richness is an increasing, non-linear function of abundance, models are needed to extrapolate or rarefy the number of samples or individuals to a common sample size (see Figure 5.2.1).

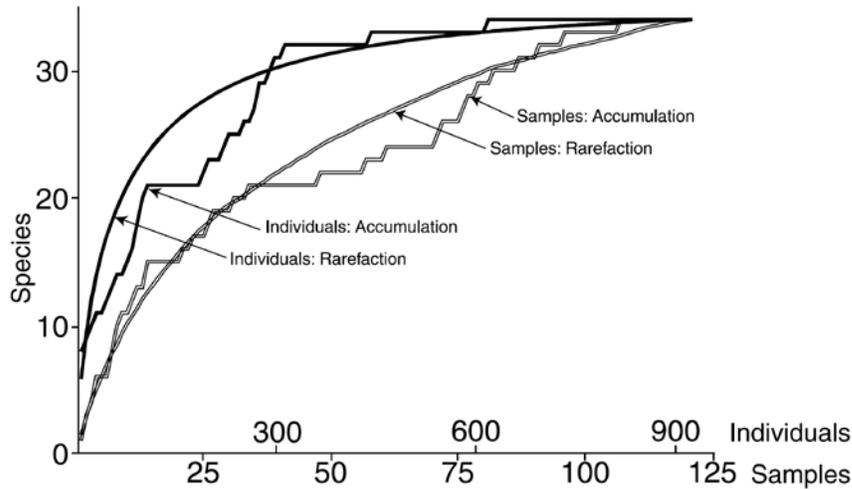


Figure 5.2.1. Sample- and individual-based rarefaction and accumulation curves. The smooth curves represent the statistical expectation of the number of species encountered based on cumulated number of collected individuals or samples (solid lines). The jagged curves represent a single ordering of the individuals or samples as they are successively pooled. Reproduced from Gotelli and Colwell (2001).

Classical rarefaction or extrapolation of samples is, however, subject to a number of factors influencing the result. Apart from differences in the local species pool generating a need for additional samples in high diversity areas, intra- and interspecific correlations and differences in aggregation patterns can violate prerequisite assumptions and change the number of samples or individuals required to produce an unbiased result. The shape of the species accumulation curve (the SAC) will depend both on the species abundance distribution (the SAD) and the species area relationship (the SAR). In addition, multicollinearity of environmental variables, differences in sampling efficiency/catchability, non-linear responses of species richness to different stressors, and effects of spatial scale may complicate the analysis (Gotelli *et al.* 2009).

The problems involved in rarefaction and extrapolation are particularly important when the effects of environmental or anthropogenic drivers are being analysed, because the shape of the SAR and SAD may change in response to changes in a driver. Sometimes SACs generated by samples obtained at different levels of, say organic enrichment, will not run parallel, but intersect providing a negative response to a change in the driver below the intersection, and a positive above the intersection point, leading to opposing conclusions about the effect (Figure 5.2.2). But even when the SACs do not intersect the relative difference between the curves may change as a function of the number of individuals or samples collected, giving rise to scale dependent conclusions generating inflated confidence limits in meta-analyses and eventually leading to misinformed conclusions about the impact of the driver.

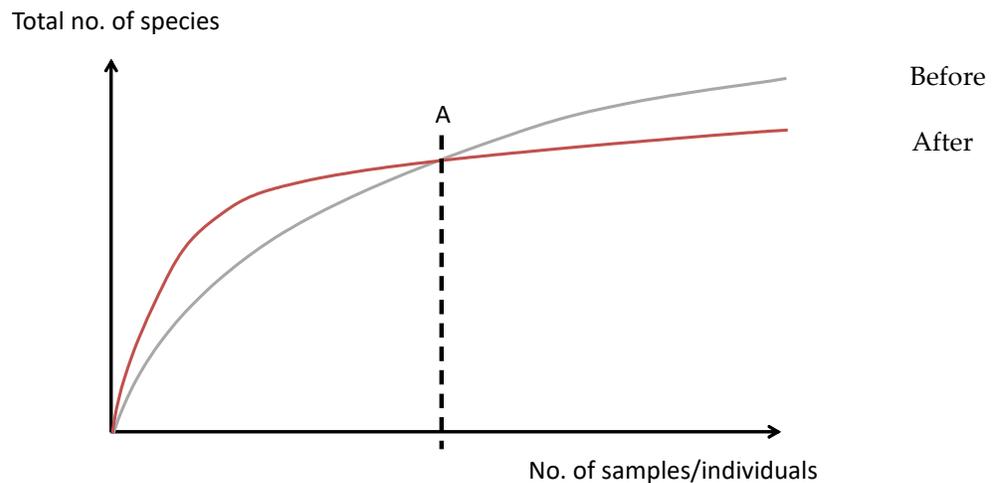


Figure 5.2.2. Species accumulation curves before and after the impact of an anthropogenic driver. At a sampling effort below the intersection A the community affected by the driver appears to be more species rich than before, above A the opposite is true. The different SACs may result from changes in aggregation, with less aggregation before than after the impact.

Chase and Knight (2013) summarize the problems related to scale dependency caused by sampling units that differ in size and spatial extent and argue that many of the problems related to identifying the direction and magnitude of biodiversity responses may be seen as resulting from the using scale-dependent methods and procedures. Except when the SACs are parallel, effect sizes will be scale dependent, creating ambiguity in estimated effect sizes.

Despite the importance of correcting for differences in spatial scale, sampling effort, and overall abundance, biodiversity samples are often not adjusted in a way that will make them comparable and allow conclusions to be drawn about changes in richness, evenness and other biodiversity measures. For the reasons given above, even samples that are standardized by area or individuals cannot generally be used without further consideration. In this section we will try to summarize the problems and their possible solution by reviewing the literature, providing examples of how the problems may influence the results, and give recommendations on how to best to analyse the impacts of environmental and anthropogenic drivers on marine biodiversity.

5.2.2 Species area relationship (SAR) and species accumulation curve (SAC)

Species Area Relationship (SAR)

The Species-Area Relationship (SAR) has been proposed as one of the ecological generalizations (Arrhenius, 1921; Rosenzweig 1995) that can be used to study species abundance and diversity both in time and space. More specifically SARs can be used to quantify changes in species diversity of fish assemblages in relation to human pressure through fishery and thus provide sensitive community-level indicators (Tittensor *et al.* 2007; Novaglio *et al.* 2016).

From theory, SAR (Arrhenius, 1921) can be described using the power function:

$$S = c \cdot A^z \quad (\text{eq. 1})$$

where S is the number of species; A is the area; and c and z are constants.

Species Accumulation Curve (SAC)

Curve-fitting methods have often been employed to obtain extrapolated richness estimates that would be expected if sample size were increased by a given amount (Colwell and Coddington 1994). For example, by re-arranging (eq. 1) the species accumulation curve (SAC) can be described by the exponential function as:

$$\log(S) = \log(c) + z \cdot \log(A) \quad (\text{eq. 2})$$

$$S = c + z \cdot \log(A) \quad (\text{eq. 3})$$

where the slope parameter (z), provides information on the rate of species accumulation as the area sampled increases (SAC). Since z is related to multiple community properties, such as species richness, species abundance, and spatial distribution, it can be used to characterize communities, and provide information on community status (Tittensor *et al.* 2007, Novaglio *et al.* 2016).

The species accumulation curve (SAC) derived from the SAR distributions using for example the exponential function: $S = c + z \cdot \log(A)$ can be used to quantify changes in species diversity in relation to fishing exploitation and its related effects on ecosystem structure, habitat, and biodiversity (Worm *et al.* 2006; Coleman & Williams 2002).

With respect to the relationship between SAC and sample size, Sanders (1968) was the first to point out that benthic samples containing different numbers of individuals need to be corrected to an equal size, before they can be compared. Traditionally this is done by 'rarefying' or interpolating the largest sample down to the size of the smallest by resampling the individuals (Magurran 2004) and methods such as the Coleman Curve or formulas have been developed to do exactly that (e.g. Hurlbert 1971). Therefore, rarefaction does not make use of the full data set.

In contrast, non-parametric estimators exist that extrapolate from the data to find the 'true' number of species, including Chao1, Chao2, abundance-based coverage estimator (ACE), incidence-based coverage estimator (ICE), and the jackknife (Gotelli and Chao 2013). These operators are based on the concept that rare species carry the most information about the number of undetected species or in other words, by using the number of rare species that are found in a sample the likelihood of other undiscovered species can be calculated.

For example, the Chao1 estimator uses the occurrence of singletons and doubletons, the number of species recorded only once or twice in a sample, respectively, to estimate the total number of individuals in the samples (Chao, 1984):

$$S_{\text{Chao1}} = S_{\text{obs}} + f_1^2 / (2f_2) \quad (\text{eq. 4})$$

where S_{obs} is the number of species in the sample, f_1 is the number of singletons and f_2 is the number of doubletons. Therefore, as soon as all species have been uncovered at least three times, it is unlikely that more species remain undetected. Chao1 works with abun-

dance data while the Chao2 estimator can be applied to incidence data. It is similar to Chao1 but incorporates a sample-size (R) correction factor $(R-1)/R$ (Chao, 1987):

$$S_{\text{Chao2}} = S_{\text{obs}} + [(R-1)/R] Q_1^2 / (2Q_2) \text{ (eq. 5)}$$

However, traditional rarefaction or extrapolation to equal-sized samples tends to misrepresent the difference in richness of the communities being compared. This is because a sample of a given size may be sufficient to fully characterize the community with the lower diversity, but insufficient to characterize the richer community. Thus, traditional methods may bias the estimate of the difference between communities of different richness.

Chao and Jost (2012) suggested to estimate the number of samples needed to remove this bias by using coverage-based sampling. In coverage-based sampling the equal-size criteria are replaced by an equal coverage criterion, where equal coverage signifies that the same proportion of the individuals per area have been sampled. They suggest a unified coverage-based sampling curve that integrates rarefaction with extrapolation by a mixture of rarefaction of the largest samples and extrapolation of the smaller ones, comparing samples of equal coverage.

5.2.3 Case studies

WGBIODIV investigated the robustness of diversity estimators with respect to sampling intensity in two case studies at differing levels of aggregation:

Hill Numbers calculated at survey level using bias correction methods based on species accumulation rates (Chao and Jost 2012; 2015)

Hill Numbers calculated for selected quarter-ICES rectangles with high sampling frequency

For both studies data were extracted from the quality assured data product created for the OSPAR Intermediate Assessment (Moriarty *et al.* 2017). Species Richness is given by Hill Numbers with $q = 0$, Shannon diversity (the exponential of Shannon entropy) at $q = 1$ and Simpson diversity (the inverse of Simpson concentration) at $q = 2$.

Case study 1: Species Accumulation Curves

We applied the approaches of Chao and Jost (2015) to determine the continuous diversity profile of Hill numbers (N_0 , N_1) annually for 19 surveys across the north east Atlantic. The techniques of Chao and Jost (2012; 2015) include methods to estimate diversity profiles from either presence-absence data (incidence) or the more traditional approaches based on abundance data. Both options were trialled with the fisheries survey data and their performance compared. Data were analysed using the “Species Prediction And Diversity Estimation” R package “SpadeR” ().

Species abundance data (catch per unit swept area, CPUA, in tonnes per km²) by haul were averaged first within appropriate sampling units (ICES rectangles for Greater North Sea surveys and by depth bands for Celtic Seas surveys) and these strata-based estimates were then summed over the whole survey area to determine annual estimated abundance. Abundance data were fourth root transformed to reduce variability in the data to the great range in catch rates and then rounded to the nearest integer since the methods trialled require count data. Note that surveys typically use subsampling methods to raise

data up to the total catch within a haul so pure count data is rarely available for any fisheries survey. Even if count data were available, variation in haul sampling volume is never fixed due to difficulties of sampling at sea and dependence on natural conditions (e.g. tide, current strength etc.). So count data, once raised to total catch, require standardisation (often by haul duration or more appropriately by estimated swept area) to take into account the different conditions between hauls.

Given that fisheries surveys typically cover great areas, individual trawls may sample multiple habitats. Because many species of fish are often found in shoals or schools it is relatively rare to find singletons and doubletons in survey data (unless a particularly odd species is caught such as a basking shark and such catches are usually excluded from the dataset due to lack of reliability of the data). Since the methods of Chao and Jost (2015) require catches of rare species (singletons and doubletons) to provide estimates of coverage, we also reran the abundance-based analyses using CPUA data scaled by the minimum CPUA value per year to make sure that singletons are present in the data.

Results

In general fisheries surveys are poor in detecting rare species, so the observed species richness will be under sampled. Estimates of expected species richness and diversity using three methods (incidence, abundance and scaled abundance) are shown for three surveys below (Figures 5.2.3 and 5.2.4) North Sea IBTS in quarter 1 (GNSIntOT1), West of Scotland GOV survey in quarter 3 (WAScoOT3), EVHOE south of 48°N in quarter 4 (CSBBFraOT4).

When species richness is estimated from incidence data only (occurrence of species by sampling unit) estimates are highly variable (Figure 5.2.3, left column). The Chao1 bias-corrected estimates of species richness based on abundance data (rounded 4th root transformed CPUA data, Figure 5.2.3, middle column) were unable to estimate coverage correctly due to the lack of singletons and doubletons. Once the abundance data were scaled (right column) so that the least abundant species was specified as a singleton, the Chao1 bias-corrected estimates of richness mirrored the estimates based on incidence data (Figure 5.2.3, left column). Overall, the estimates of species richness are highly variable and thus unlikely to mirror the change in richness within the ecosystem from year to year. The Chao1 bias-corrected estimates of richness demonstrate the total community sampled by the surveys is highly variable due to the nature of the sampling and hence we conclude that richness estimated in this way are not suitable for biodiversity indicators. Shannon diversity is far less variable than richness and here the Chao and Jost estimates make very little difference to the estimates directly from the data (Figure 5.2.4). Notably the scaled abundance data do inflate the diversity estimates since rarity has been artificially increased.

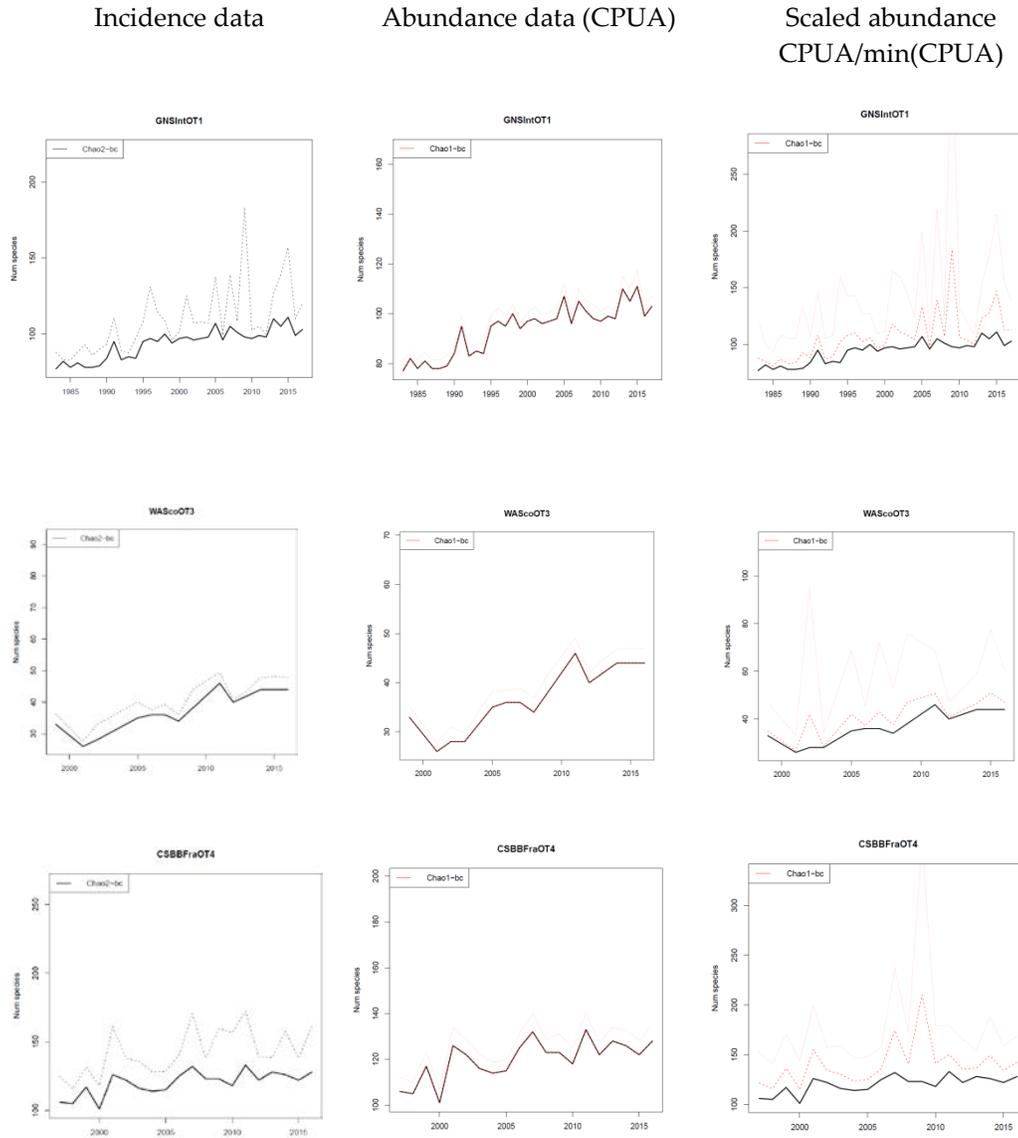


Figure 5.2.3. Annual estimates of species richness using incidence data i.e. occurrence of species (left column, empirical and bias corrected estimate of richness), CPUA data (central column, showing empirical and bias corrected estimate of richness in red with upper 95% CI bound in dashed red line) and scaled CPUA data (CPUA/min(CPUA)),right column) for selected surveys: GNSIntOT1=IBTS Q1, WAScoOT3=West of Scotland GOV survey, CSBBFraOT4=EVHOE south of 48°N. The bias corrected estimator of species richness is based on the methodology of Chao and Jost (2012).

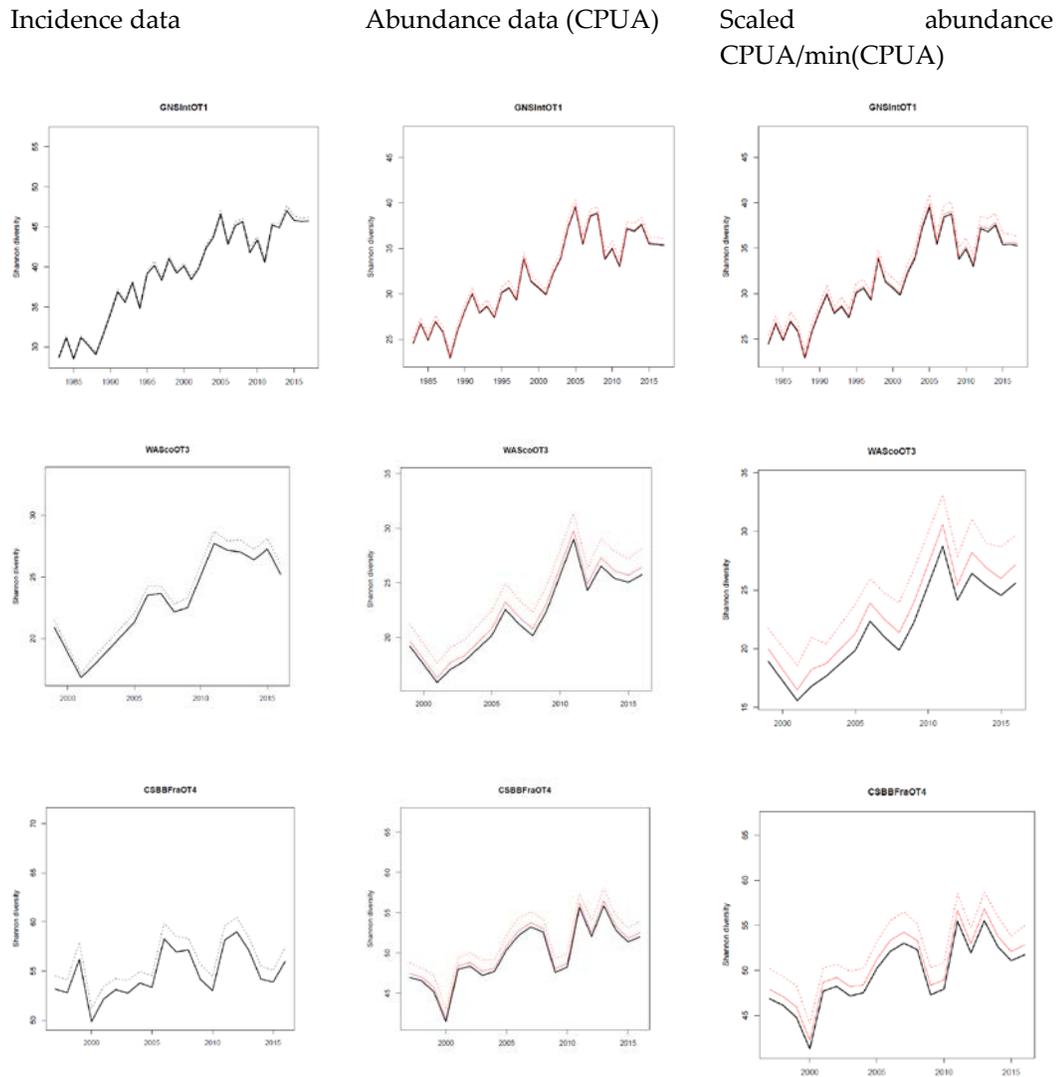


Figure 5.2.4. Annual estimates of species diversity (Shannon diversity, Hill N1) using incidence data i.e. occurrence of species (left column, empirical and bias corrected estimate of richness, dotted line), species diversity using CPUA data (central column, showing empirical and bias corrected estimate in red with upper 95% CI bound in dashed red line) and species diversity using CPUA/min(CPUA) data (right column) for selected surveys: GNSIntOT1=IBTS Q1, WAScoOT3=West of Scotland GOV survey, CSBBFraOT4=EVHOE south of 48°N. The empirical estimate of Shannon diversity is the maximum likelihood estimate and the bias corrected estimate uses the methodology of Chao *et al.* 2013.

The application of SAR and SAC theory needs further consideration for the development of ecosystem status indicators that can be used to quantify biodiversity loss across ecosystems. The approach by Chai and Jost (2015) to estimate Hill-numbers are not suitable for trawl survey data. To develop reliable spatial estimates of diversity metrics from the observed data, further study is required to determine the minimum sampling size that is acceptable (see study 2 below).

Case study 2: Species Area Relationships obtained from fisheries surveys

The abundance of rare objects in nature (rare species, rare length classes within a species) is difficult to assess with precision (Magurran 1988; 2004; Colwell *et al.*, 2004; Colwell and Coddington, 1994), and this is especially so when, as with groundfish surveys, the fraction of the environment actually sampled is extremely small. The 'standard survey area' of the GNSIntOT1 survey covers 176 ICES statistical rectangles (0.5° latitude by 1.0° longitude), amounting to a total sea area of 556 187 km². Over the course of the 35y survey time-series, the total area swept between the wings of the GOV trawl in any one year varied between 20.7 km² and 41.1 km², equivalent to between 0.004% and 0.007% of the total sea area. The problem is then further confounded by the fact that the catchability (trawl catch efficiency) of many species and of many size classes of particular species, in the GOV trawl is low. Hence, the proportion of all sampled individual fish present in the environment may actually be an order of magnitude (or more) smaller than the estimates based on the fraction of the area sampled might suggest (Fraser *et al.*, 2007; Walker *et al.*, 2017).

Comparison of abundance estimates derived from small-scale high-intensity (1 trawl sample of ½h duration per 15 km²) and large-scale low-intensity (1 trawl sample of ½h duration per 1500 km²) suggest that data collected from at least 4 to 6 trawl samples per ICES statistical rectangle (approximately 3090 km²) require to be aggregated before adequate estimates of local species abundance at length are obtained (Stransky, 1998). Early studies of North Sea fish species diversity suggested that aggregation of data from 10 1h trawl samples or 20 ½h trawl samples may be necessary to derive reliable estimates of local (point or alpha diversity) species richness and species evenness (Greenstreet and Hall, 1996; Greenstreet *et al.*, 1999; Greenstreet and Piet, 2008). Aggregation of samples to derive reliable parameter estimates can be carried out either across space, to derive a value for the aggregate area for a given point in time (e.g. year), or across time, to derive a value for a given location (e.g. ICES statistical rectangle) for the aggregate period. However, particularly in respect of species diversity metrics, both types of sample aggregation present potential difficulties in that the desired estimates of alpha diversity become confounded by incorporation of elements of beta diversity, i.e. the turnover in species composition associated with habitat variability in space and environmental change over time (Velland, 2001, Korhonen *et al.*, 2010; Magurran *et al.*, 2010; Soinen, 2010; Kraft *et al.*, 2010). Greenstreet and Piet (2008) examined the spatial aspect of this issue explicitly and found that spatial beta diversity started to inflate estimates of alpha diversity in any given year when samples collected at distances of greater than 50km to 60km apart were aggregated.

To construct annual maps of spatial variation in fish species- and community-level metrics, the 176 ICES statistical rectangles that constituted the GNSIntOT1 'standard survey area' were each subdivided into four quadrants. The intention here was to derive maps with a spatial resolution (pixel size) that was much closer to the resolution used in determining the shape and size of Marine Protected Areas (MPAs) than the ICES statistical rectangle resolution that underpins the survey design. Ideally, even finer resolution would be useful in this respect, but the splitting of ICES statistical rectangles into four quadrants was considered to be the finest resolution that could be supported by the available data. Quadrants with a sea area of <76% were excluded assuming these to be too close to land to be properly part of the survey area.

Annual estimates of the density-at-length of each species present in each quadrant in each year were determined as the mean species density-at-length calculated across all hauls collected in each year within a 60km radius of the central point of each quadrant. Over the full 35y time-series of the GNSIntOT1 survey, species density-at-length estimates could be determined in this way for a total of 677 quadrants. Figure 5.2.5 shows variation in the number of individual trawl samples contributing to each quadrant's species density-at-length estimate in each year. Frequently, no estimates could be determined for some quadrants because no trawl samples were collected within a 60km distance from the central location of these quadrants in a particular year. This occurred most commonly in quadrants located at the edges of the study area. Other edge effects were also apparent; often the number of hauls available within 60km of edge quadrant central points was relatively low. Some species- and community-level metrics are influenced by variation in sampling effort. Where this is the case, these plots of annual quadrant sampling effort can provide an indication of confidence in particular quadrant metric values. Where the number of trawl samples available is particularly low, estimates of species richness are likely to be biased downwards, while estimates of species diversity, dominance and evenness are likely to be either extreme positive or extreme negative outliers (Greenstreet and Piet, 2008).

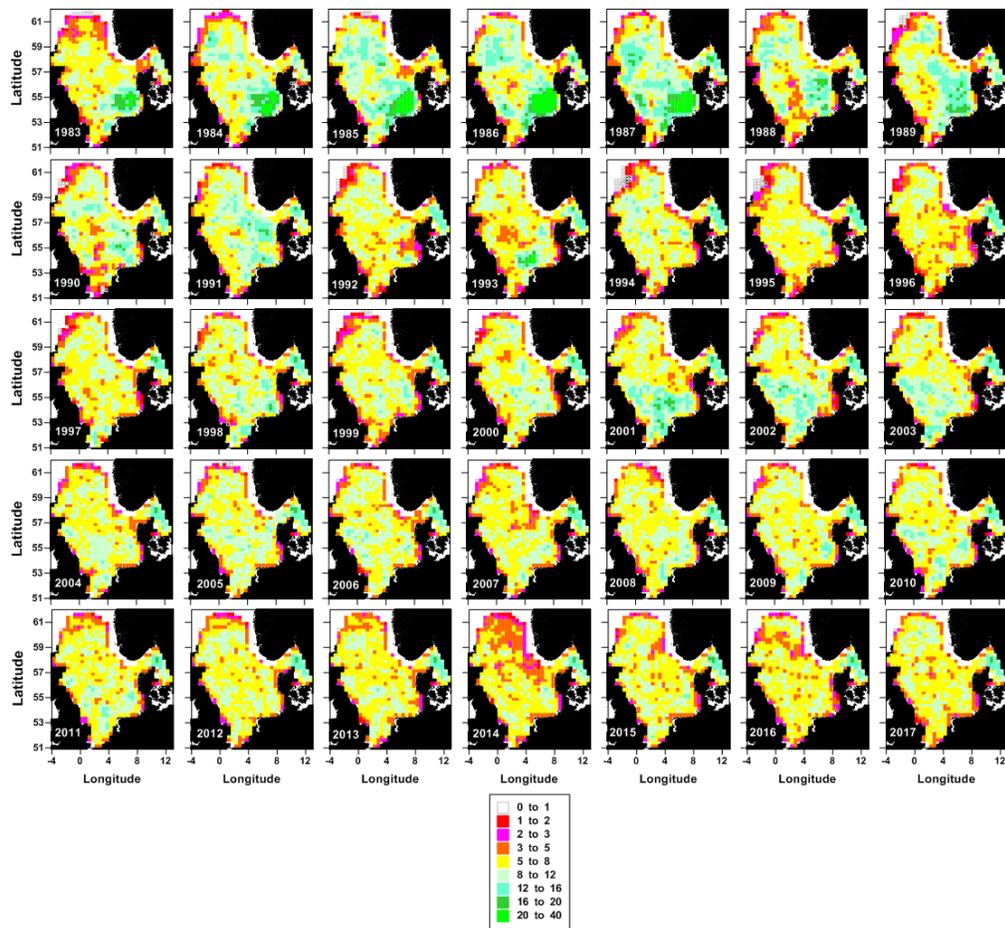


Figure 5.2.5. Spatial variation in the number of hauls within a 60km radius of the central point of each ICES statistical rectangle quadrant (0.25° latitude by 0.5° longitude) in each year.

Assessing levels of sampling effort necessary to derive reliable indicator estimates

This poses the obvious question – what level of sampling is necessary in order to derive reliable estimates of species richness and evenness? Four spatially contiguous quadrants have 17 or more trawl samples collected in four consecutive years (1984 to 1987) in the southern North Sea of the Dutch coast (Figure 5.2.6). The effect of sample size on resulting quadrant species count and diversity metric estimates was examined for each quadrant in each year. Samples were aggregated starting from the sample collected closest to each quadrant’s central point location and then adding samples increasingly further away from the central point. Figures 5.2.7–5.2.10 show the resulting metric accumulation curves for species count S (Hill’s N_0), Hill’s N_1 diversity, Hill’s N_2 dominance and Pielou’s J evenness metrics. In accordance with previous studies of North Sea demersal fish species diversity (Greenstreet and Piet, 2008), Gleason ($Y = b \times \ln(X) + a$) semi-log relationships generally provided the closer fit than the more usual Arrhenius ($Y = a * X^b$) power function suggesting a degree of limitation in the available species pool (Table 5.2.1). Examination of these curves suggests that a sampling effort level of 10 or more ½h trawl samples is necessary before estimates of S , N_1 and N_2 even start to approach satu-

ration. This is not the case for J , which approaches saturation at sampling effort levels closer to five ½h trawl samples.

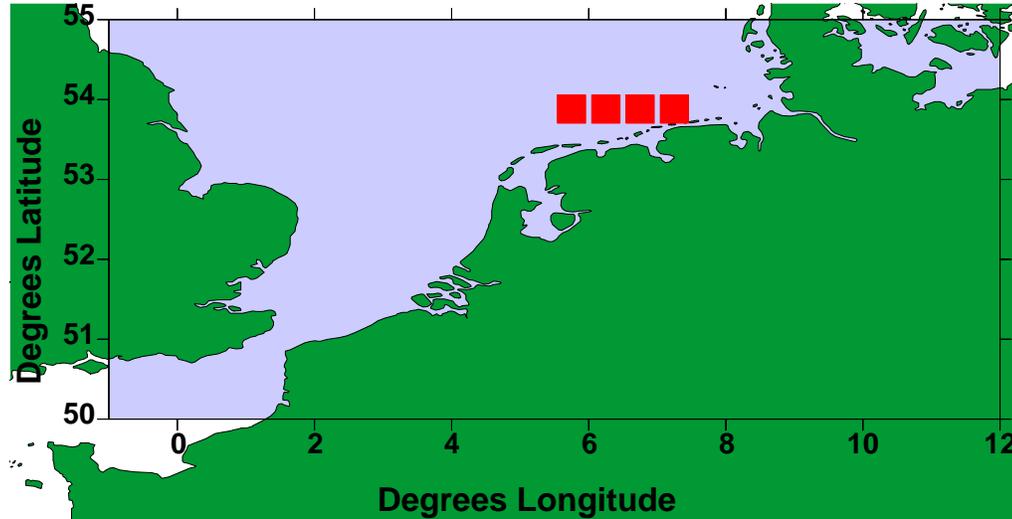


Figure 5.2.6. Location of four ICES statistical rectangle quadrants intensively sampled in four consecutive years.

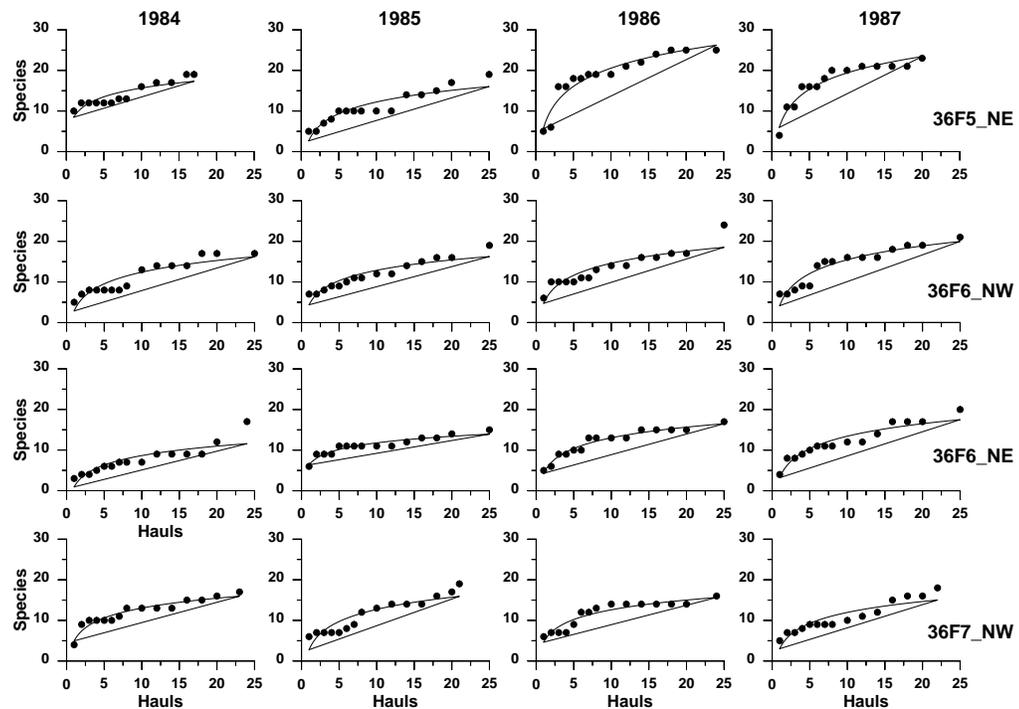


Figure 5.2.7. Effect of increase in the number of trawl samples on species count S estimates in four contiguous ICES statistical rectangle quadrants in four consecutive years.

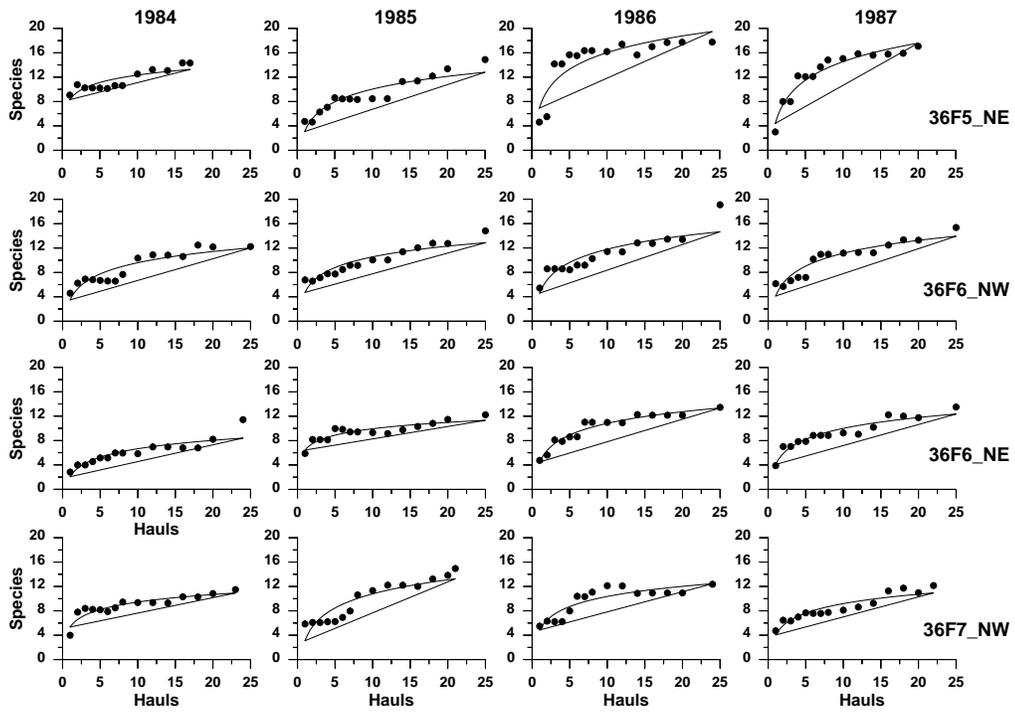


Figure 5.2.8. Effect of increase in the number of trawl samples on Hill's N1 species diversity estimates in four contiguous ICES statistical rectangle quadrants in four consecutive years.

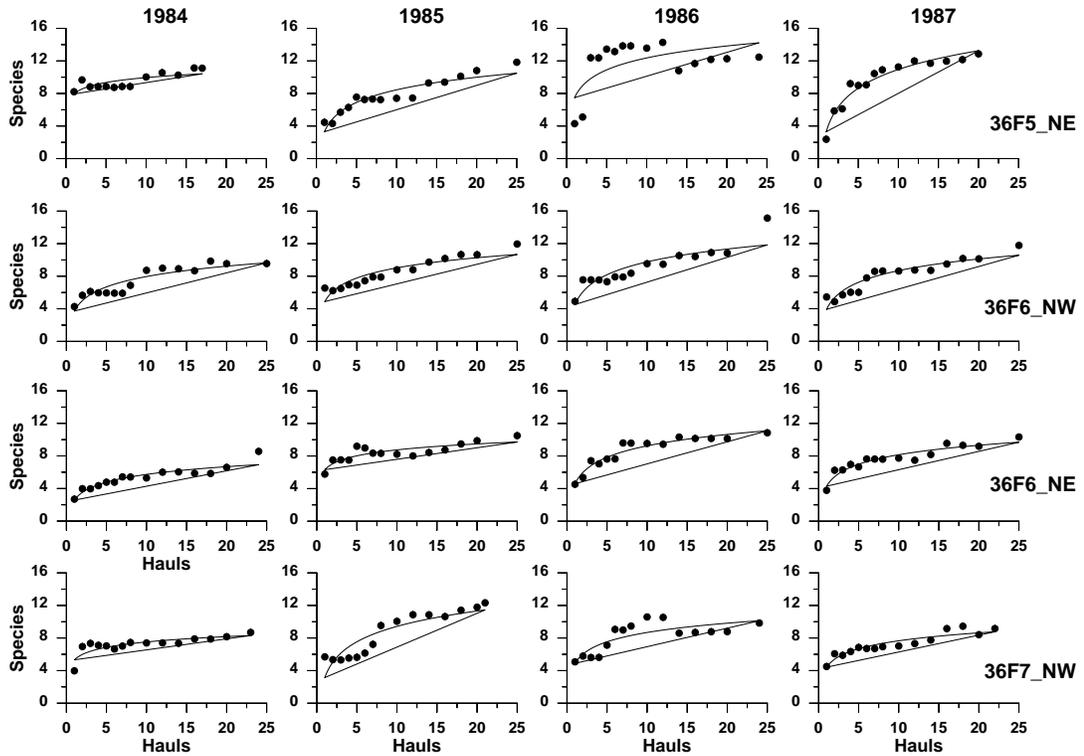


Figure 5.2.9. Effect of increase in the number of trawl samples on Hill's N2 diversity estimates in four contiguous ICES statistical rectangle quadrants in four consecutive years.

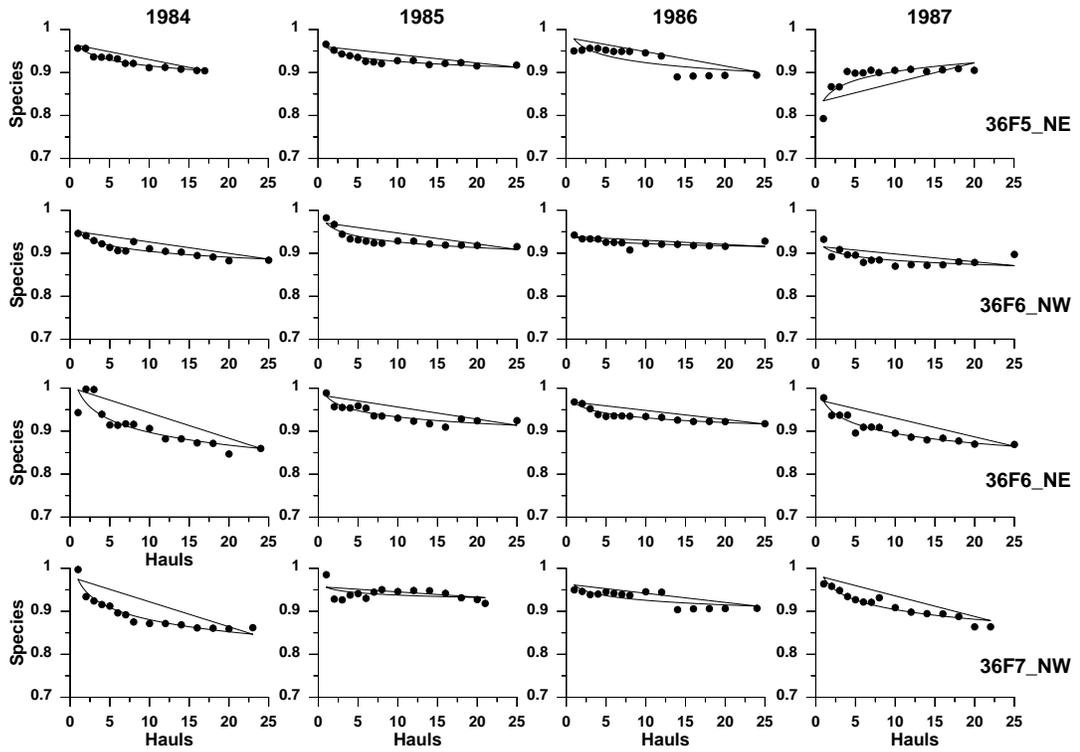


Figure 5.2.10. Effect of increase in the number of trawl samples on Pielou’s species evenness J estimates in four contiguous ICES statistical rectangle quadrants in four consecutive years.

Table 5.2.1 Gleason semi-log fit parameter values for plots shown in Figures 5.2.7 to 5.2.10.

Quadrant\Year	S		$N1$		$N2$		J	
	b	a	b	a	b	a	b	a
36F5_NE1984	3.1252	8.471	1.7522	8.301	0.8711	7.947	0.02059	0.9630
36F5_NE1985	4.1517	2.655	3.0232	3.069	2.2328	3.305	-0.01468	0.9596
36F5_NE1986	6.4805	5.629	3.9557	6.905	2.1177	7.489	-0.02410	0.9784
36F5_NE1987	5.8212	5.974	4.3943	4.412	3.3213	3.299	0.02956	0.8340
36F6_NE1984	4.1449	2.869	2.6526	3.486	1.8401	7.722	-0.01976	0.9502
36F6_NE1985	3.6905	4.375	2.5351	4.721	1.7974	4.886	-0.01901	0.9702
36F6_NE1986	4.2860	4.721	3.1320	4.583	2.2751	4.510	-0.00709	0.9386
36F6_NE1987	4.9133	4.137	3.0506	4.109	2.0645	3.922	-0.01353	0.9145
36F6_NW1984	3.3488	0.932	1.9930	2.060	1.3703	2.572	-0.04287	0.9958
36F6_NW1985	2.3478	6.385	1.5183	6.444	1.0560	6.321	-0.02110	0.9818
36F6_NW1986	3.7914	4.307	2.7533	4.489	2.0225	4.589	-0.01574	0.9674
36F6_NW1987	4.4403	3.213	2.5717	4.083	1.6778	4.293	-0.03258	0.9699
36F7_NW1984	3.5066	4.961	1.7801	5.343	0.9519	5.323	-0.04069	0.9743
36F7_NW1985	4.3193	2.771	3.3406	3.089	2.7290	3.142	-0.00798	0.9561
36F7_NW1986	3.4625	4.639	2.3925	4.848	1.6479	4.883	-0.01543	0.9612
36F7_NW1987	3.8803	3.029	2.2330	4.037	1.4075	4.416	-0.03257	0.9790

However, it is not necessary to derive indicator values that approach “actual” species richness and evenness. All that is required is that the indicator values obtained provide sufficient indication of “actual” values such that they can reliably detect differences in species richness and evenness in space and time. The Gleason models in Table 5.2.1 were used to extrapolate estimates of S , $N1$, $N2$ and J for each quadrant and year that would be obtained by 50 ½h trawl samples and these were assumed to represent “actual” values at each location and time. The effectiveness of estimates derived from lower levels of sampling effort was first examined theoretically by regressing (with intercept set to zero) Gleason model estimates derived from smaller aggregations of samples (as the dependent variable) in each quadrant and year with “actual” estimates (as the independent variable); (Figure 5.2.11). The regression essentially examines how well Gleason estimates derived from a range between 5 and 40½h trawl samples predict the “actual” values derived from Gleason models based on 50 ½h trawl samples. The regression slope therefore provides an estimate of the factor difference between “actual” values and values derived from lower sampling effort levels. For example, “actual” species richness estimates of S derived from 50 ½h trawl samples were likely to be on average 1.86 times higher than estimates derived from Gleason models based on five ½h trawl samples. Despite this though, Gleason model estimates of S derived at sampling effort levels of five ½h trawl samples still differentiated between the 16 quadrant/year observations reasonably well giving an R^2 value of 0.986 for the regression between dependent and actual metric values. The observed estimates of S , $N1$, $N2$ and J obtained from various levels of sample aggregation also compared well with the “actual” i.e. theoretical value estimate. Examining the same example, the Gleason estimate of “actual” S were on average 1.93 times higher than estimates derived from aggregation of five ½h trawl samples and the R^2 value of 0.953 suggested that the estimates of S determined from aggregation of five ½h trawl samples were capable of differentiating between the 16 quadrant/year observations with a high level of success. This analysis suggests that in studies of spatial and temporal variation in groundfish species diversity, it is necessary to aggregate at least five ½h trawl samples in order to derive reliable metric estimates.

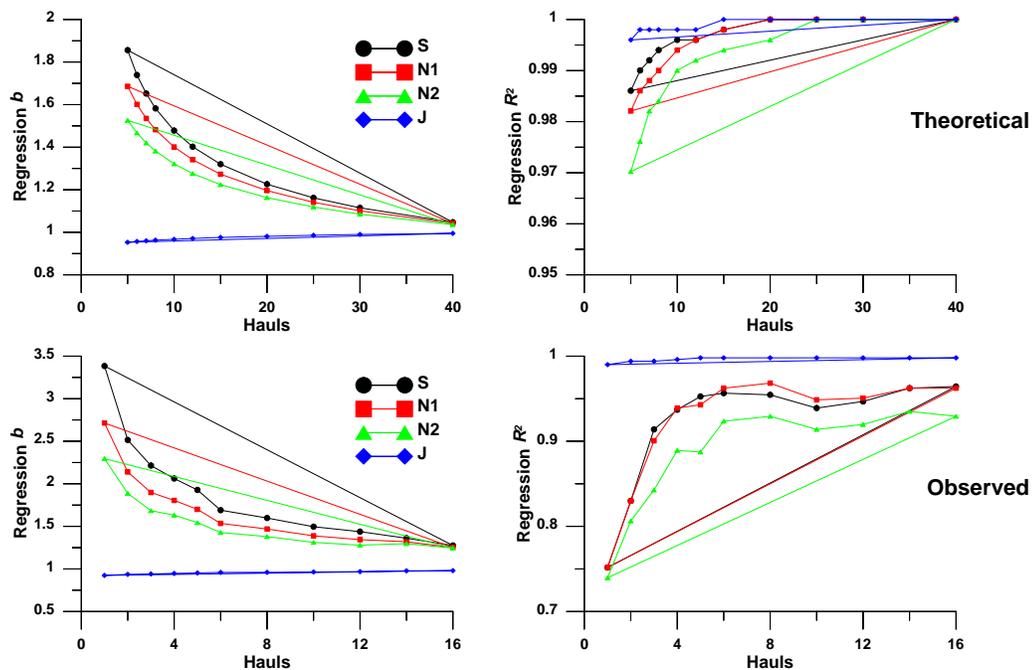


Figure 5.2.11. Comparisons of estimates of species richness (S), and species diversity (N1, N2, J) derived from Gleason semi-log functions fitted to sample aggregation data obtained from four ICES statistical rectangle quadrants in four consecutive years when at least 17 ½h trawl samples were collected on each occasion. In each case, the extrapolation of the Gleason models to an aggregated sample of 50 ½h trawl samples was assumed to represent the “actual” values present in each quadrant and year. The “theoretical” comparison compares the Gleason model estimates from Table 5.2.1 with these “actual” estimates and the “observed” comparison compares the observed metric estimates derived from various levels of sample aggregation with the “actual” estimates.

Case study 3: On the spatial resolution of monitoring for calculating plankton indicators

The spatial representation of two OSPAR pelagic habitats indicators (PH1 and PH2) was explored within the EcApRHA project (Ostle *et al.*, 2017) by combining classical monitoring data acquired by hydrological bottle/plankton net at a fixed station (L4, western Channel long-term series Plymouth) and data gathered by the Continuous Plankton Recording (SAHFOS), both followed by microscopic observation and discrimination of phytoplankton and zooplankton species. CPR samples were averaged from extending zones around the fixed station L4 (Figure 1).

Amongst the indicators discussed by Ostle *et al.* (2017), PH1 features a "Plankton Index" (PI, derived from the PCI of Tett *et al.*, 2008 and Tett *et al.* 2013) of lifeform pairs (e.g. dinoflagellates vs. diatoms). The PI has been developed to track changes in the state of the plankton in marine waters over time by: (i) the grouping of plankton species into functional types or lifeforms; (ii) the display of changes in the abundance of each of these lifeforms using a state-space approach; (iii) by calculating a Plankton Index (PI) quantifying possible changes in the state of the plankton relative to baseline or starting conditions; and (iv) relating trends in the PI to trends in human pressures and climate change indices (when possible). Ecologically-relevant lifeform pairs must be selected according to the marine system considered, in order to make lifeforms ratio changes indicative of changes in: hydrological/nutrient conditions and pathways to different primary produc-

ers (i.e. changes in pigmentary dinoflagellates and diatoms, changes in microphytoplankton and nano- and pico-phytoplankton, changes in pelagic and thycopelagic phytoplankton); the transfer of energy from primary to secondary producers (changes in phytoplankton and zooplankton); the pathway of energy flow and top predators (changes in gelatinous zooplankton and fish larvae); benthic/pelagic coupling (changes in holoplankton i.e. fully planktonic) vs. meroplankton, only part of the lifecycle is planktonic, the remainder is benthic) (Gowen *et al.*, 2011). For the spatial analysis between L4 and CPR data, the monthly lifeforms were calculated for CPR data within each spatial range shown in Figure 5.2.12 and the annual PI time-series were compared with L4 for the period 2004 to 2014 (with a reference period of 2004 to 2008; Ostle *et al.*, 2017).

The PH2 indicator is based on the consideration of the temporal anomalies in the distribution of total phytoplankton biomass and the total copepod abundance. For comparison with the CPR dataset, total phytoplankton biomass was inferred from the Phytoplankton Colour Index (PCI), and total copepod abundance was determined in a similar way to the above but using CPR counts. The anomalies were calculated following the methodology used in the time-series R package Pastecs by F. Ibanez *et al.* (2006) by de-seasonalising the monthly data using the monthly mean cycle, and calculating the annual average anomaly from the monthly anomalies.

Seasonal and annual anomalies of PH2 were addressed and results at fixed station monitoring were compared to continuous acquiring samples with CPR, whereas PH1 was considered at an annual basis.

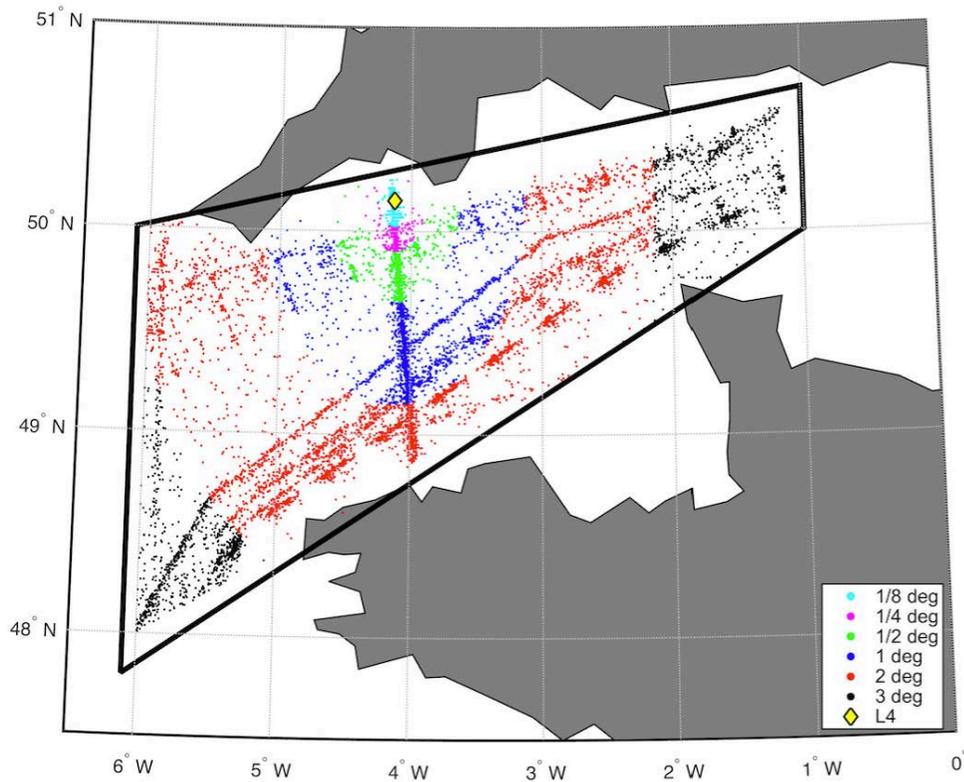


Figure 5.2.12. Map of the sample locations of the EcAprHA study (Ostle *et al.*, 2017). The coloured dots represent CPR samples from extending ranges around L4, which is represented as a yellow diamond. Cyan dots = CPR samples within 1/8 degree of L4. Pink dots = CPR samples within 1/4 degree of L4. Green dots = CPR samples within 1/2 degree of L4. Dark blue dots = CPR samples within 1 degree of L4. Red dots = CPR samples within 2 degrees of L4. Black dots = CPR samples within 3 degrees of L4. The black polygon represents extent of study area used.

According to the results of this study on the western Channel (Ostle *et al.*, 2017) spatial variation in plankton abundance, as a consequence of its patchy distribution, limits the ability to merge differing datasets inter-annually, e.g. for PH2 calculation on total copepod abundance (Figure 5.2.13). However, on mean seasonal and long-term time scales both datasets show strong similarities on PH2 seasonal cycle for total copepod abundance (Figure 5.2.14) and PH1 inter-annual trends on "auto and mixotrophic dinoflagellates/diatoms" life form (Figure 5.2.15).

From further analyses, it is apparent that there is a balance between having an adequate number of samples for aggregation within a close enough region around the fixed-point station to make a meaningful comparison with the results of a single station alone. Long-term and seasonal trends can be compared and inferred using both sets of data, however fixed-point and transect data have differing complementary information for time-series investigations.

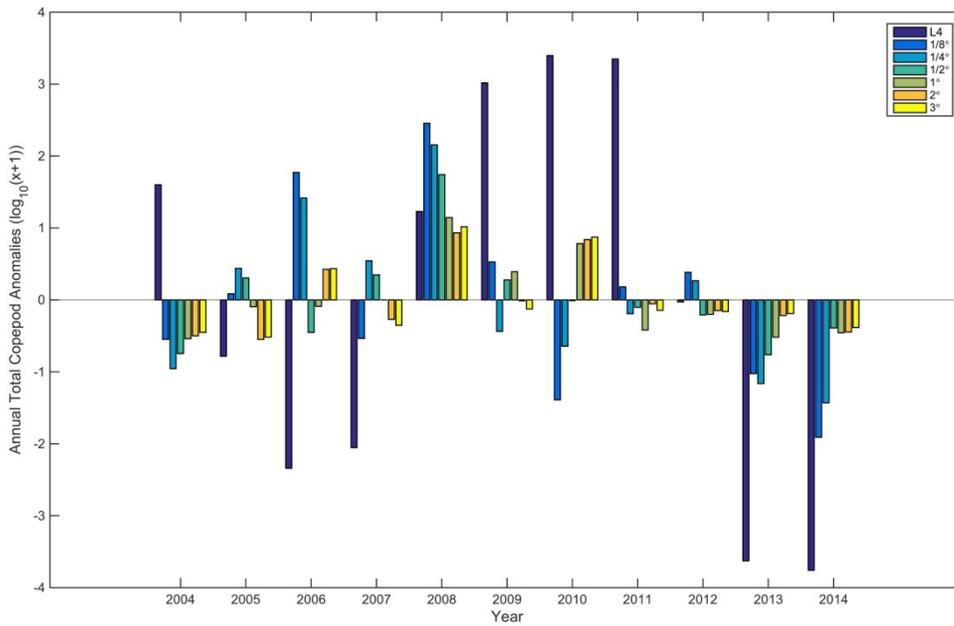


Figure 5.2.13. Annual anomalies in total copepod abundance for L4, and CPR data within the spatial ranges of L4 shown in Figure 1 for the years 2004 to 2014. (Ostle *et al.*, 2017).

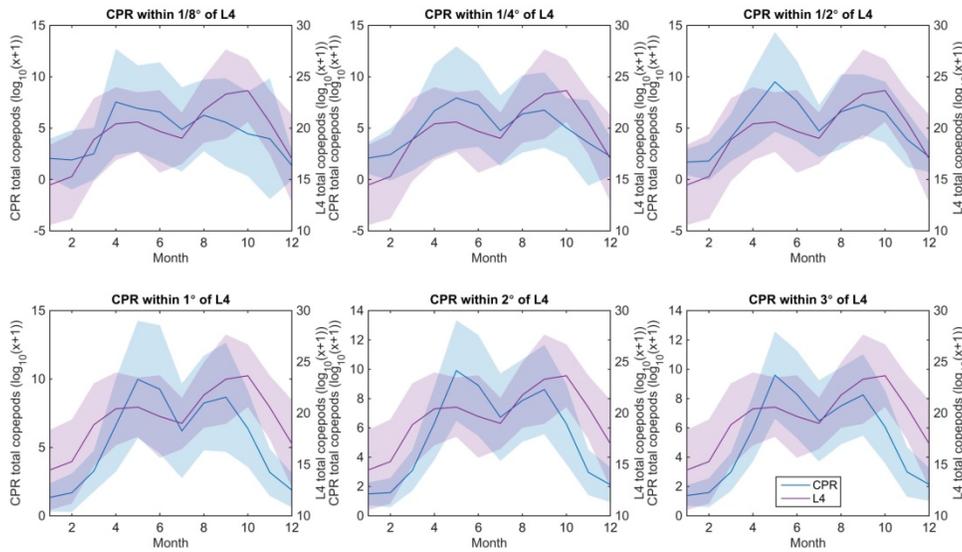


Figure 5.2.14. Mean monthly seasonal cycle of total copepod abundance. Purple = L4, blue = CPR samples at the different spatial ranges around L4 shown in Figure 1. The bounded area around the lines represents the standard deviation of the mean. (Ostle *et al.*, 2017).

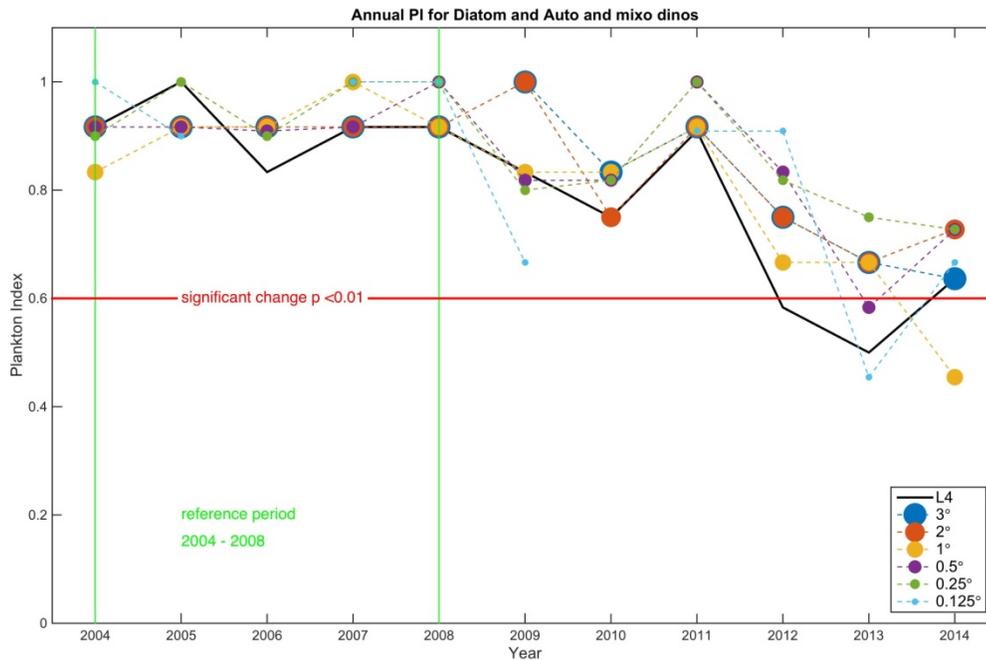


Figure 5.2.15. Annual PI values for the lifeform pair diatoms vs. auto and mixo-trophic dinoflagellates for L4 and CPR data within the spatial ranges shown in Figure X.12. Black line = L4, dark blue circles = 3 degrees, red circles = 2 degrees, orange circles = 1 degree, purple circles = 1/2 degrees, green circles = 1/4 degrees, cyan circles = 1/8 degrees. The reference period used for this analysis was 2004 to 2008, and falls between the green lines. A significant change ($p < 0.01$) in the annual PI is characterised by a drop in the PI value to below 0.6, this threshold is represented as the red line. (Ostle *et al.*, 2017).

Indeed, further investigation is needed to determine the optimum sample size which is ecologically representative of a water mass. It is likely that this will vary between regions and indicators used, as physical dynamics play a strong role in structuring the spatial distribution of plankton abundance, biomass, productivity and diversity. Therefore, this shall have to be carried out on a site-by-site basis, however, a protocol for this determination of optimum sample size could be developed by analysing the stability of the average at the lowest resolution using a boot-strapping technique. Pelagic habitats or ecohydrodynamic (EHD) areas (as defined for the North Sea by van Leeuwen *et al.*, 2015) should be used to test how representative samples from the same EHD are (no matter which sampling approach is applied). A state-space version of a multivariate autoregressive model that accounts for observation error was proposed by Holmes *et al.* (2010). If there should be a potential to adapt the plankton index state-space approach to account for observational error, perhaps a framework to incorporate spatial variation could also be incorporated in to the methodology for PH1 (to be further investigated).

For better defining the spatial representation of a fixed-point station, more information is required on the hydrological features surrounding a station and further analyses of individual taxa and their range in relation to those environmental conditions are necessary. On the other hand, there is a need for further investigation on how to adapt existing pelagic habitats indicators to data provided by seasonal cruises covering wide spatial scales, thus, higher spatial resolution but lower temporal frequency than CPR (which

does not cover all marine systems) and stations monitoring (which are most of the time limited to coastal areas).

Case study 4: Relationship between sampling effort and species number – case study soft bottom benthos

Soft bottom fauna has been used for environmental monitoring for over a century. This ecosystem compartment constitutes one of the biological quality elements in monitoring according to the Water Framework Directive, and is used in the overall classification of water bodies (aiming at a state of at least “good”). In such monitoring, it is highly important that the methodology in sampling and processing is standardized between the institutions involved. The first step in this chain is the sampling itself, which is usually conducted with a van Veen grab (0,1 m²), but also other sampling gears like box-cores may be used. These may sample the fauna differently, particularly regarding surface-crawlers and deep-living fauna. Another aspect is how many replicates are sampled. The common soft bottom standard ISO 16665 is highly specific in several aspects of the sampling and processing, but does not explicitly recommend the number of replicates to be sampled. Though the number of replicates may vary depending on the purpose of the monitoring and the habitat under subject, it should in many cases be more standardized.

With regards to costs and-benefits, it is important to balance sampling effort vs. accuracy of diversity metrics. Quite a lot of work has been done to look into which species are lost when reducing the sampling effort. However, less is known on whether they are functionally important – i.e. what ecological information is lost? This is particularly important to have information on as biological trait analysis (BTA) of the faunal communities is increasingly used. In one of the few studies conducted on this topic, Ellingsen *et al.* (2007) concluded that rare species are important to functional biodiversity of soft bottom benthos.

Here a very small case study based on data from two stations in the Norwegian Coastal Monitoring Program is presented (Figure 5.2.16), where we compare a sampling effort of 0,4 m² (i.e. 4 van Veen grabs) with 0,8 m² (i.e. 8 grabs) for one shallow (50 m) and one deep (360 m) station. The species accumulation curve (Figure 5.2.17) did not show any signs of flattening out on any of the stations.

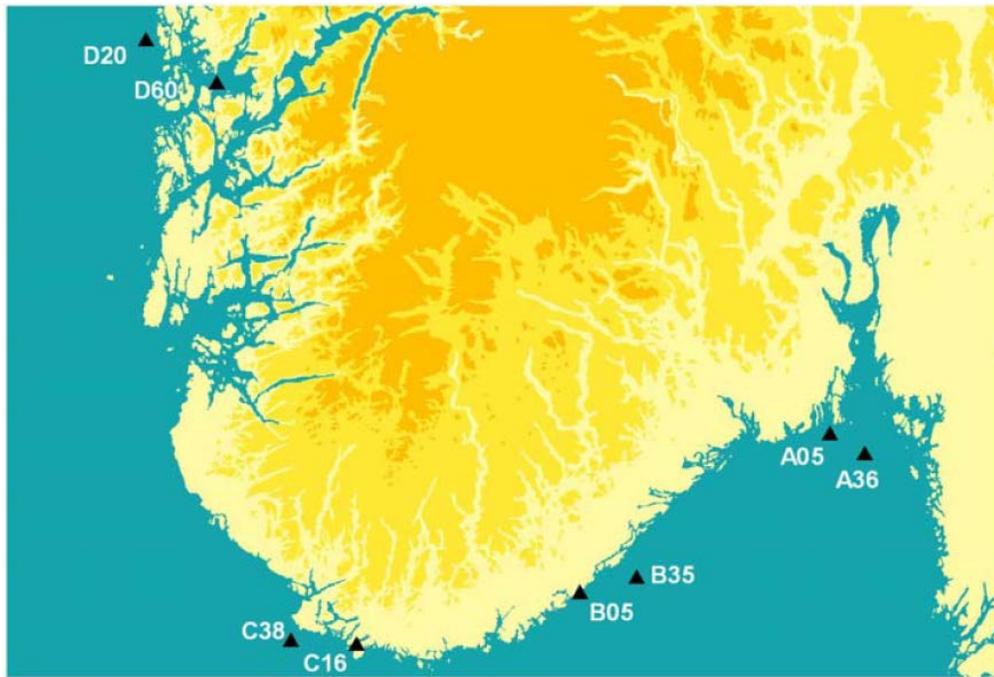


Figure 5.2.16. Soft bottom stations within the Norwegian Coastal Monitoring Programme. The stations were positioned in three regions: the outer Oslofjord (A), the southeast coast (B), and the southwest coast (C). Depth is the number multiplied with 10.

Station A05

Station A36

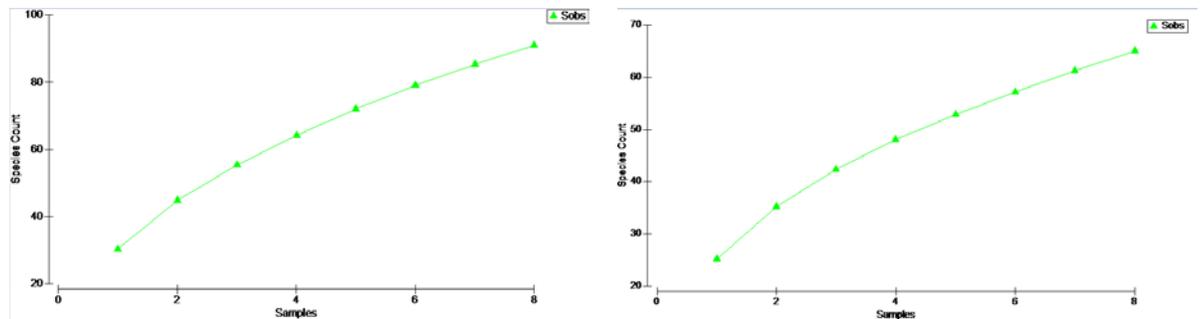


Figure 5.2.17. Species accumulation curve for the two Norwegian Coastal Monitoring Stations A05 (50 m depth) and A36 (360 m depth), 2009.

The number of species, number of individuals as well as some selected diversity indices are presented in Table 5.2.2. This clearly shows a strong increase in number of species when the sampling effort is increased from four to eight samples. At the same time it should be noted that the “extra species” recorded in eight vs. four samples all had low density; most were in fact singletons, with one major exception of *Thyasira sp.* having a number of 143 in the first four replicates at station A36.

Table 5.2.2. Overview of species richness (S), total abundance N, Simpson-Index (d), Pielou's evenness (J)', Hulbert-Index (ES100) and Shannon-Diversity (H) for either eight or four samples of the stations A05 and A36 in 2009.

	S	N	d	J'	ES(100)	H'(log2)
A05 (8 samples)	91	606	14,05	0,82	39,90	5,37
A05 (4 samples)	55	271	9,64	0,87	36,82	5,02
A36 (8 samples)	65	2023	8,41	0,62	19,68	3,75
A36 (4 samples)	50	1213	6,90	0,64	18,39	3,59

The second aim of this small case study was to look into the properties of the “lost” species regarding their traits. If the “lost” species have a disproportional influence on the community, this should be taken into account in the evaluation of the sampling effort. Below is an overview of species present in the four first, but not four last replicates, which then potentially are lost with fewer replicates (Table 5.2.3). NIVA has a soft bottom database of approximately 1300 species, and for half of these traits have been described (Table 5.2.4). The traits were then extracted for the taxa in Table 5.2.3 and associated with scores (Table 5.2.5). In this very brief analysis, it was not possible to see any strong tendency of a particular size categories of the “lost” species. Notably, several of the “lost” species at the shallow station (A05) live in a semi-permanent tube. At the deep station (A36) most of the “lost” species were burrowers, and some surface crawlers. At this station, the “lost” species were also characterized by none degree of attachment, which was not true for A05. Neither stations showed any particular affinity towards one particular feeding mode. To conclude, the “lost” species show a variety of traits, and there was also an interesting difference between the two sampling depths regarding mobility pattern of the “lost” species.

Table 5.2.3. Overview of “lost” species present in the first four, but not four last replicates, for one shallow and one deep station in the Coastal Monitoring Program, 2009.

A05

(50m)

Nephtys ciliata	0	0	0	4	0	0	0	0
Glycera rouxii	1	0	0	2	0	0	0	0
Exogone sp	0	0	0	2	0	0	0	0
Trichobranchus roseus	0	0	0	2	0	0	0	0
Thyasira equalis	0	0	1	1	0	0	0	0
Corbula gibba	2	0	0	0	0	0	0	0
Holothuroidea indet	2	0	0	0	0	0	0	0
Phyllodoce rosea	0	0	0	1	0	0	0	0
Podarkeopsis helgolandica	0	0	1	0	0	0	0	0
Ceratocephale loveni	0	0	0	1	0	0	0	0
Orbinia sertulata	0	0	0	1	0	0	0	0
Prionospio cirrifera	0	0	0	1	0	0	0	0
Chaetopterus variopedatus	0	0	0	1	0	0	0	0
Hydrobiidae indet	1	0	0	0	0	0	0	0
Aporrhais pespelecani	0	1	0	0	0	0	0	0
Arctica islandica	0	1	0	0	0	0	0	0

A36 (360 m)

Thyasira sp	27	0	112	3	0	0	0	0
Scalibregma inflatum	1	1	1	0	0	0	0	0
Harmothoe sp	0	0	1	1	0	0	0	0
Nephtys paradoxa	1	0	0	1	0	0	0	0
Levinsenia gracilis	0	0	0	2	0	0	0	0
Anobothrus gracilis	0	0	0	2	0	0	0	0
Mugga wahrbergi	0	2	0	0	0	0	0	0
Harpinia cf. pectinata	0	0	0	2	0	0	0	0
Irregularia indet	2	0	0	0	0	0	0	0
Aphrodita aculeata	1	0	0	0	0	0	0	0
Nephtys ciliata	0	0	1	0	0	0	0	0
Nephtys pulchra	0	0	1	0	0	0	0	0
Paraonis fulgens	0	0	0	1	0	0	0	0
Ophelina acuminata	0	0	1	0	0	0	0	0
Oligochaeta indet	0	0	1	0	0	0	0	0
Bivalvia indet	0	0	0	1	0	0	0	0
Yoldiella lucida	0	0	0	1	0	0	0	0
Delectopecten vitreus	0	0	0	1	0	0	0	0
Thyasira flexuosa	0	1	0	0	0	0	0	0
Erichthonius abditus	0	1	0	0	0	0	0	0
Amphilepis norvegica	1	0	0	0	0	0	0	0
Vermiformis indet	1	0	0	0	0	0	0	0

Table 5.2.4. List of traits and modalities in the NIVA traits database. The traits are scored according to the fuzzy coding method, i.e. allowing several traits pr. taxa (0 = no affinity, 1 = low importance, 2 = moderately high importance, 3 = dominant).

FCODE	TEXT	TYPE
AH1	Adult_life_habitat1	Sessile
AH2	Adult_life_habitat2	Tube (permanent)
AH3	Adult_life_habitat3	Tube (semi-permanent)
AH4	Adult_life_habitat4	Burrower
AH5	Adult_life_habitat5	Surface crawler
AH6	Adult_life_habitat6	Swimmer
AM1	Releative_adult_mobility1	None
AM2	Releative_adult_mobility2	Low
AM3	Releative_adult_mobility3	Medium
AM4	Releative_adult_mobility4	High
BF1	Body_form1	Short cylindric
BF2	Body_form2	Flattened dorsally
BF3	Body_form3	Flattened laterally
BF4	Body_form4	Ball shaped
BF5	Body_form5	Long thin, treadlike
BF6	Body_form6	Irregular
DA1	Degree_of_attachment1	None
DA2	Degree_of_attachment2	Temporary
DA3	Degree_of_attachment3	Permanent
FD1	Faecal_deposition1	Sediment surface
FD2	Faecal_deposition2	Subsurface 0–5cm

FD3	Faecal_deposition3	Deep subsurface >5cm
FH1	Feeding1	Suspension/filter
FH2	Feeding2	
FH3	Feeding3	Surface deposit feeder, SDF
FH4	Feeding4	Subsurface deposit feeder, DF
FH5	Feeding5	Dissolved matter/symbionts
FH6	Feeding6	Large detrius/scrapper/grazer (inkl. tidl. FH2)
FH7	Feeding7	Scavenge
FH8	Feeding8	Carnivore/omnivore
FH9	Feeding9	Parasite/commensal
LD1	Life_duration1	< 1 year
LD2	Life_duration2	1–5 year
LD3	Life_duration3	>5 year
LT1	Larval_type1	Planktotroph (feeding larvae)
LT2	Larval_type2	Lecitotroph (non-feeding larvae)
NS1	Normal_adult_size1	<5mm
NS2	Normal_adult_size2	5mm-1cm
NS3	Normal_adult_size3	1–3cm
NS4	Normal_adult_size4	3–6cm
NS5	Normal_adult_size5	6–10cm
NS6	Normal_adult_size6	>10cm
NY1	Number_of_reproductive_cycles_per_year1	< 1
NY2	Number_of_reproductive_cycles_per_year2	1
NY3	Number_of_reproductive_cycles_per_year3	2 or more
RP1	Reproductive_period1	December-February
RP2	Reproductive_period2	March-May
RP3	Reproductive_period3	June-August
RP4	Reproductive_period4	September-November
RP5	Reproductive_period5	no particular season
RT1	Reproductive_technique1	asexual (budding)
RT2	Reproductive_technique2	broadcast spawner
RT3	Reproductive_technique3	Demersal eggs
RT4	Reproductive_technique4	Brooder, viviparous
SD1	Sediment_dwelling_depth1	0 cm (surface)
SD2	Sediment_dwelling_depth2	0–1cm
SD3	Sediment_dwelling_depth3	1–5cm
SD4	Sediment_dwelling_depth4	5–15cm
SD5	Sediment_dwelling_depth5	>15 cm

- Coleman, F. C., Williams, S. L. 2002. Overexploiting marine ecosystem engineers: potential consequences for biodiversity. *Trends in Ecology & Evolution*, **17**, 40–44.
- Colwell R.K. and Coddington J.A. 1994. Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions of the Royal Society of London Series B*, **345**: 101–118.
- Colwell R.K., Mao C.X. and Chank J. 2004) Interpolating, extrapolating, and comparing incidence-based species accumulation curves. *Ecology*, **85**: 2717–2727.
- Ellingsen, K.E., Hewitt, J.E., Thus, h, SF., 2007. Rare species, habitat diversity and functional r The species–accumulation curve and estimation of species richness. *Journal of Sea Research* **58**: 291–301
- Fraser H.M., Greenstreet S.P.R. and Piet G.J. 2007. Taking account of catchability in groundfish survey trawls: implications for estimating demersal fish biomass. *ICES Journal of Marine Science*, **64**:1800–1819.
- Gotelli, N. J., & Colwell, R. K. (2001). Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology letters*, **4**(4), 379–391.
- Gotelli, N.J. *et al.* (2009). Patterns and causes of species richness: a general simulation model for macroecology. *Ecology Letters*, **12**, 873–886.
- Gotelli, N.J, & Chao, A. (2013). Measuring and estimating species richness, species diversity, and biotic similarity from sampling data. In: Levin, S.A., Ed., *Encyclopedia of Biodiversity*, Vol. 5, Academic Press, Waltham, MA: 195–211
- Greenstreet, S.P.R. and Hall, S.J. 1996. Fishing and the ground-fish assemblage structure in the north-western North Sea: an analysis of long-term and spatial trends. *Journal of Animal Ecology*, **65**: 577–598.
- Greenstreet, S. P. R. and Piet, G. J. 2008. Assessing the sampling effort required to estimate α species diversity in groundfish assemblages of the North Sea. *Marine Ecology Progress Series*, **364**: 181–197.
- Greenstreet, S.P.R., Spence, F.E. and McMillan, J.A. 1999. Fishing effects in northeast Atlantic shelf seas: patterns in fishing effort, diversity and community structure. V. Changes in structure of the North Sea groundfish assemblage between 1925 and 1996. *Fisheries Research*, **40**: 153–183.
- Gowen, R.J., McQuatters-Gollop, A., Tett, P., Best, M., Bresnan, E., Castellani, C., Cook, K., Forster, R., Scherer, C. and Mckinney, A., (2011). *The Development of UK Pelagic (Plankton) Indicators and Targets for the MSFD*. Advice to Defra, Belfast, UK, 41 pp.
- Hurlbert, S. H. (1971). The non-concept of species diversity: a critique and alternative parameters. *Ecology*, **52**(4), 577–586.
- Ibanez F., Grosjean P. & Etienne M. (2006) CRAN Pastecs: Package for Analysis of Space-Time Ecological Series. <http://www.sciviews.org/pastecs>
- Magurran AE 1988. *Ecological diversity and its measurement*. Chapman & Hall, London.
- Magurran, A. E. 2004. *Measuring biological diversity*. Blackwell Oxford UK.
- Moriarty, M., Greenstreet, S.P.R. and Rasmussen, J. (2017) Derivation of Groundfish Survey Monitoring and Assessment Data Product for the Northeast Atlantic Area. *Scottish Marine and Freshwater Science* Vol 8 no 16, 240pp. DOI: 10.7489/1984–1
- Novaglio, C., Ferretti, F., Smith, A. D. M. & Frusher, S. 2016. Species – area relationships as indicators of human impacts on demersal fish communities. *Diversity and Distribution*, **22**, 1186–1198.

- Ostle, C., Artigas, F., Atkinson, A., Aubert, A., Budria, A., Graham, G., Helaouët, P., Johns, D., Padegimas, B., Rombouts, I., Widdicombe, C. and McQuatters-Gollop, A., (2017). Spatial Representativity of Plankton Indicators EcApRHA Deliverable WP1.3. OSPAR, London, UK, 28.
- Rosenzweig, M. L. 1995. *Species diversity in space and time*. Cambridge University Press
- Sanders, H. L. (1968). Marine benthic diversity: a comparative study. *The American Naturalist*, 102(925), 243–282.
- Stansky, C. 1998. A comparison of biomass estimates obtained from small and large-scale ground-fish surveys in the North Sea. Unpublished MSc Thesis, University of Aberdeen, Scotland, UK.
- Tett P, Carreira C, Mills DK, van Leeuwen S, Foden J, Bresnan E, Gowen RJ. 2008. Use of a phytoplankton community index to assess the health of coastal waters. *ICES Journal of Marine Science*, 65: 1475–1482.
- Tett P, Gowen RJ, Painting SJ, Elliott M, Forster R, Mills DK, Bresnan E, Capuzzo E, Fernandes TF, Foden J, Geider RJ, Gilpin LC, Huxham M, McQuatters-Gollop AL, Malcolm SJ, Saux-Picart S, Platt T, Racault MF, Sathyendranath S, van der Molen J, Wilkinson M. 2013. Framework for understanding marine ecosystem health. *Marine Ecology Progress Series*, 494: 1–27.
- Tittensor, D. P., Micheli, F., Nyström, M. & Worm, B. 2007. Human impacts on the species–area relationship in reef fish assemblages. *Ecological Letters* 10, 760–772.
- van Leeuwen, S., Tett, P., Mills, D. and Molen, J.v.d., (2015). Stratified and nonstratified areas in the North Sea: Long-term variability and biological and policy implications. *Journal of Geophysical Research Oceans*, 120: 4670-4686.
- Walker, N. D., Maxwell, D. L., Le Quesne, W. J. F., and Jennings, S. 2017. Estimating efficiency of survey and commercial trawl gears from comparisons of catch-ratios. *ICES Journal of Marine Science: Journal du Conseil*, 74: 1448–1457.
- Worm, B., Barbier, E. B., Beaumont, N., Duffy, J. E., Folke, C., Halpern, B. S., *et al.* 2006. Impacts of biodiversity loss on ocean ecosystem services. *Science*, 314, 787–790.

5.3 Development of a Benthic Community Physical Disturbance Indicator

Deliverable 5. One or more operational indicators to assess biodiversity at the community and eventually the ecosystem level.

5.3.1 Introduction

Many economic sectors depend on marine ecosystems and biodiversity. The conservation, restoration, more equitable sharing of benefits and sustainable use of biodiversity is necessary to ensure that marine ecosystems continue to provide the goods and services on which human populations depend (Balmford *et al.*, 2002). Over recent decades marine ecosystems have been subject to a growing variety and extent of human pressures (Halpern *et al.* 2008), giving rise to concern over the impact on the state of marine ecosystems (Worm *et al.* 2006). Consequently, the need for holistic ecosystem based management (EBM) of marine natural resources has become increasingly apparent (Jennings 2004). EBM is shifting the focus towards more comprehensive decision-making processes by recognising ecological systems as a rich mixture of interacting elements and by acknowledging their social and economic features (Gregory *et al.* 2013).

Biodiversity is inherently a multi-dimensional concept, spanning genes and species, functional forms, adaptations, habitats and ecosystems, as well as the variability within and between them. All these dimensions of biodiversity are tightly interconnected, affecting the state, stability, and productivity of marine ecosystem as well as ecosystem services (Laurila-Pant *et al.* 2015). By altering populations of target species, fishing has effects on the structure and functioning of ecosystems. The activity of fishing also has effects on ecosystem structure and functioning. Over the last century, the use of bottom trawls and other mobile fishing gears have greatly increased in intensity and geographic extent (Watson *et al.* 2006). Through contact with the seafloor, these gears can damage and modify benthic habitats and their associated benthic communities (Kaiser *et al.* 2000), kill or remove target and non-target species (Kaiser *et al.* 2006) and truncate age and size distributions (Jørgensen *et al.* 2007).

Many studies designed to assess the effects of fishing on benthic communities focus on taxonomic approaches, implying that all species are equally important. However, species do not contribute equally to ecosystem processes and differ markedly in their responses to changing environments. This recognition has led to the exploration of other components of biodiversity, notably ecologically important traits (Dencker *et al.* 2017). Traits are measurable functional and structural attributes of species. These attributes can be morphological (e.g. size and body shape), physiological (e.g. metabolic pathways or growth related) or behavioural (e.g. diurnal migration, feeding patterns). Combinations of traits can describe the ecological niche of species and determine their response to environmental gradients and perturbations (Dencker *et al.* 2017 and refs therein).

A complex interplay of evolutionary and environmental factors underpin the occurrence of species in different places. These factors determine species' fitness. Thus, explaining species distributions in response to specific environmental conditions, natural or human-mediated, should benefit from consideration of traits related to fitness, reflecting growth, survival and reproduction (Lavorel *et al.*, 2008). Although biological traits are increasingly being used in marine community ecology, their potential for reflecting functional responses has rarely been considered in the framework of ecological indicator development (Beauchard *et al.*, 2017).

5.3.2 Policy rationale

The Convention of Biological Diversity (CBD <http://www.biodiv.org>; <https://www.cbd.int/doc/legal/cbd-en.pdf>) is the first treaty in international law to emphasise the vital importance of biodiversity conservation. More recently, the European Union (EU) has emphasised the importance of biodiversity. The EU Biodiversity Strategy (<http://ec.europa.eu/environment/nature/info/pubs/docs/brochures/2020%20Biod%20brochure%20final%20lowres.pdf>) has become an important policy driver and biodiversity is one of the descriptors of Good Environmental Status in the Marine Strategy Framework Directive (MSFD; <https://www.eea.europa.eu/policy-documents/2008-56-ec>).

The MSFD represents the first instance of the application of EBM that covers all human activities in, and pressures on, the marine environment, addresses all major aspects of marine ecosystems, and which is fully underpinned by binding legislation. The MSFD requires Member States (MSs) to co-operate, through the auspices of the Regional Seas Conventions, to achieve good environmental status (GES) at the regional seas scale by 2020. Where possible, MSs and non-EU countries bordering shared regional seas are re-

quested to harmonise their management strategies and use the same indicators to monitor change in environmental status (ICES 2015).

Paragraph 25 of the MSFD preamble states that MSs “[...] should determine for their marine waters a set of characteristics for good environmental status”, which it defines as “*the environmental status of marine waters where [...] the structure, functions and processes of the constituent marine ecosystems, together with the associated physiographic, geographic, geological and climatic factors, allow those ecosystems to function fully and to maintain their resilience to human-induced environmental change. Marine species and habitats are protected, human-induced decline of biodiversity is prevented and diverse biological components function in balance*”, i.e. the quality and occurrence of habitats and the distribution and abundance of species need to be at, or achieve, good environmental status.

For the seafloor including its benthic community and how this is impacted by the pressure “physical disturbance” two criteria apply according to the MSFD:

- D6C2: Spatial extent and distribution of physical disturbance pressures on the seafloor.
- D6C3: Spatial extent of each habitat type which is adversely affected, through change in its biotic and abiotic structure and its functions (e.g. through changes in species composition and their relative abundance, absence of particularly sensitive or fragile species or species providing a key function, size structure of species), by physical disturbance.

Clearly, these two criteria are (or should be) interrelated through a pressure-state relationship in that the pressures, i.e. D6C2 “Spatial extent and distribution of physical disturbance pressures on the seafloor”, determines the state, i.e. D6C3 “Spatial extent of each habitat type which is adversely affected, through changes in species composition and their relative abundance”.

5.3.3 Scientific rationale

Efforts to infer the status of seafloor communities tend to use information on the abundances and biomass of species within them. For example, the effects of physical disturbance on benthic biomass may be described using a logistic population growth model with a depletion rate due to a trawling event and a recovery rate (Pitcher *et al.*, 2016):

$$dB/dt = rB(1-B/K) - dFB$$

where dB/dt is the rate of change in abundance B in time t (years), r is the logistic recovery rate (y^{-1}), K is carrying capacity, d is depletion rate (fraction killed by a single trawl pass, specific to different gear-types) and F is trawling intensity (calculated as swept area in a grid cell divided by surface area of that grid cell, $year^{-1}$).

The change in biomass (i.e. impact) of a benthic population caused by physical disturbance (i.e. fishing) can therefore be captured by two habitat- and gear-specific parameters, the depletion rate d and the recovery rate r (Figure 5.3.1).

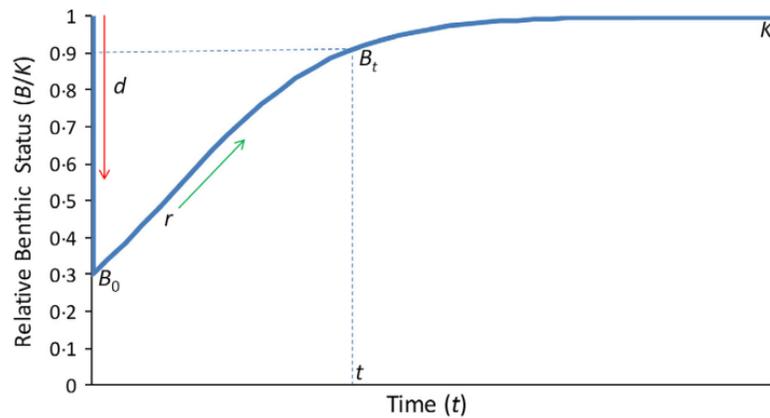


Figure 5.3.1. Conceptual diagram showing the two components of physical disturbance effects on benthic communities (biomass B): gear-specific depletion rate d and habitat-specific recovery rate r , reaching the carrying capacity K over time. Redrawn from Pitcher *et al.* (2016).

As exact depletion and recovery rates are only available for a few species, we use trait-based information to estimate the relative responsiveness of species (or at least a representative sample) in the benthic community. We specifically focus on those traits and their modalities that we expect to respond to physical disturbance. The assumption is that species that differ in those traits will be impacted differently by varying intensities of fishing pressure.

Sensitivity expresses instantaneous damage or death from physical disturbance, which results in biomass depletion. It is assumed to be shaped by three traits:

- The fragility (FR) of a species, with more fragile species being more depleted.
- The motility (MO) of a species determines its ability to avoid or escape disturbance, with more mobile species capable of avoiding disturbance, reducing their depletion.
- The vertical position of a species in the sediment (burrowing depth, BD) determines its exposure to disturbance (or predation), with deeper-burrowing species depleted less.

Thus, depletion is assumed to decrease with burrowing depth and moving speed, and to increase with body fragility.

Recovery reflects the likelihood that a population can recover from depletion. Three traits were chosen to reflect this:

- Offspring type (OT) describes the reproductive mode of a species. Species broadcasting eggs are considered to reproduce more slowly than species releasing larvae after brooding. Species which produce juveniles (direct development) after incubating their young generally recover more rapidly by ensuring a higher probability of offspring survival in situ.
- Age at sexual maturity (AM) influences recovery because a species that does not reach sexual maturity for some time has a higher probability of dying from physical disturbance before reproducing successfully. For instance, if an organism reaches AM early and is reproductively active over several years, sev-

eral cohorts can be reproductive at any one time, increasing the resilience of the total population. This life history aspect is expressed as the reproductive life span, equal to life span (LS) minus AM. Synergistically, offspring survival increases recovery potential.

- Depending on AM, the life span (LS) of a species can affect recovery because a long reproductive life span (LS minus AM) may ensure a minimum of successful reproductive events after reaching maturity in contrast to the opposite, and extreme, case of semelparity (i.e. only one reproduction before dying).

The modalities of these traits were scored to reflect increasing responsiveness to physical disturbance due to depletion and decreasing ability to recover subsequently (Table 5.3.1). Since traits were composed of different numbers of modalities (from three to four), a standardisation was applied to balance the influence of each trait in the computation of the indicator.

Table 5.3.1. Traits and their modalities used in the computation of the indicator. Scores express the degree of sensitivity to physical disturbance, except for life span, which is used in the recovery component to express reproductive life span (see section 5.3.4 below).

Trait	Modality	Raw score	Standardized score
Fragility	Robust	1	0.00
	Intermediate	2	0.50
	Fragile	3	1.00
Motility	Swimmer	1	0.00
	Crawler-Swimmer	2	0.33
	Tubicolous	3	0.67
	Attached	4	1.00
Burrowing depth	>15 cm	1	0.00
	5-15 cm	2	0.33
	0-5 cm	3	0.67
	0 cm	4	1.00
Maturity	<1 year	1	0.00
	1-3 years	2	0.50
	>3 years	3	1.00
Life span	<1 year	1	0.00
	1-3 years	2	0.33
	3-10 years	3	0.67
	>10 years	4	1.00
Offspring type	Juvenile	1	0.00
	Larva	2	0.50
	Egg	3	1.00

5.3.4 Constructing a traits-based indicator

The previously defined sensitivity and recovery components are species' functional properties that can be combined to derive three species types (resistant, resilient and vulnerable) with varying degrees of vulnerability to physical disturbance. For example, vulnerable types are characterized by high sensitivity (large instantaneous depletion) and low recoverability (long recovery time); (Table 5.3.2).

Table 5.3.2. Species typology in response to the effect of physical disturbance. Each species type is defined by a combination of functional properties (sensitivity and recoverability). Note that low sensitivity and high vulnerability is not evidenced in the “Life history approach” from ICES WKBENTH 2017 (ICES, 2017).

Functional property	Species type		
	Resistant	Resilient	Vulnerable
Sensitivity	None	Low or High	High
Recoverability	Not applicable	High	Low

Hence, the indicator incorporates a sensitivity component, taking into account the instantaneous effect of physical disturbance such as contact with a beam trawl during a haul, and a recovery component considering key aspects of population dynamics.

The sensitivity component (SE) was expressed by the synergistic effect of the BD, MO and FR:

$$SE = FR \times MO \times BD$$

The recovery component (RE) was computed as:

$$R = \frac{AM}{LS - AM + 1} \times AM \times OT$$

with 1 added to the reproductive life span to deal with cases where both would equal 1.

A short reproductive life span ($LS - AM + 1$) negatively affects recovery because the probability of a species reproducing successfully is lower. This lower probability of successful reproduction is exacerbated if maturity is reached later (high AM). A high ratio of AM upon reproductive life span thus implies lower recoverability. An extreme example is a semelparous species, that reproduces only once. As it is a ratio, however, it may take the same value for species with different ages of maturity and reproductive life spans, e.g. a species maturing at 1 year of age and dying after 2 years will give the same ratio as another one maturing at 2 years and dying after 4 years. The second species can support three reproductive cohorts (one more than the first species), but its later AM exposes it to a higher risk of dying before reproducing. To account for this, the ratio is multiplied by AM to include an appropriate penalty to recoverability for long lived organisms that reach AM late.

The combination of both SE and RE captures the effects of trawling that proceed over time. An important assumption is that resilient and vulnerable organisms can be equally sensitive (Table 5.3.2). When bottom trawling frequency is high, the most sensitive organisms from a previously undisturbed community are extirpated instantaneously. Over time, the most resistant organisms remain so that the further removal of benthic biomass progresses more and more slowly (Figure 5.3.2A). The recovery potential is also expected to decline over time, with a progressively increasing amount of biomass extirpated per unit of time. This effect would be indirect and act through age at maturity. If the trawling frequency exceeds age at maturity (e.g. age of 3 years, one trawl every year), the density of reproductively active and surviving individuals would become too low over time to ensure the replacement of individuals that were extirpated by trawling. Since larger organisms are generally the most fecund, this should also hold for total biomass density

and even biomass per individual. In the longer term, the remaining organisms are those of low SE, i.e. those benefiting from high reproductive allocation and with direct development that maximizes survival and settlement (Figure 2B).

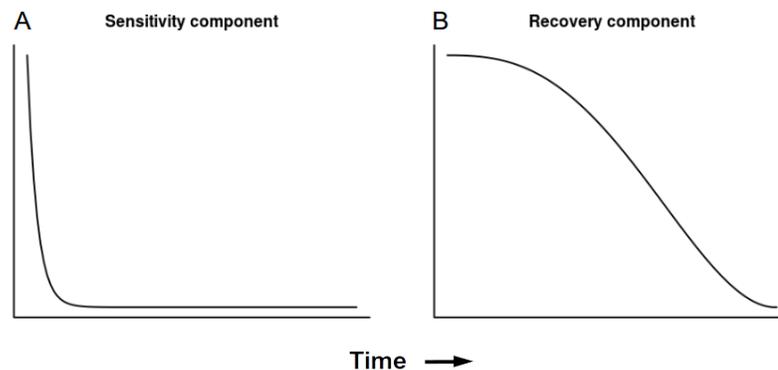


Figure 5.3.2. Response of the sensitivity component (A) and the recovery component (B) to bottom trawling over time under constant trawling pressure.

It is not clear *a priori* how the SE and RE components should be combined in a single formula to ascribe a global score of vulnerability to a community of species. If SE and RE are independent they may simply be summed. If they are interactive, they may be multiplied. For instance, if the community is dominated by highly sensitive organisms of low recoverability, an interactive effect is to be considered to explain the response of the benthic indicator. Therefore, testing the sensitivity of the individual components and their additive and interactive combinations was a necessary step in developing the benthic indicator and studying its response.

5.3.5 Application of the indicator to two different case studies

The behaviour and performance of the benthic indicator was examined using two independent case studies. For both of these, fauna could be assigned to a common set of traits and data on the intensity of bottom trawling were available. The indicator components were structured so that high scores were related negatively to beam trawling intensity. Two basic criteria were applied to validate the benthic indicator: the sign of response variation (positive or negative) and the significance of the explained variance.

5.3.5.1 Dutch EEZ

The Dutch EEZ ranges from 51.62°N to 55.47°N where 81 stations were yearly sampled between 1995 and 2012. Information on data collection and the construction of the traits data base is described in detail in the ICES WKBENTH report (ICES, 2017). Since beam trawl fishing intensity was strongly correlated to the main environmental gradient in the study area, the specific effect of beam trawling on the benthic indicator had to be disentangled from the effects of habitat variables (Legendre and Legendre, 2012). Figure 5.3.3 illustrates how the benthic indicator was computed from trait data. In this case study data set, field data include individual and biomass densities so that three types of density

where applied, including presence-absence data, as well as absolute and relative densities (% of total sample density).

In this study area, two habitats strongly contrast: a shallow sandy habitat (20 to 40 m deep) and a deep muddy one (30 to 54 m deep); (Figure 5.3.4). Hence, the indicator was tested at two different spatial scales, over the whole area and per habitat.

Fishing intensity (swept area ratio) was quantified as part of the BENTHIS project (Eigaard *et al.*, 2017) from VMS data obtained from 2001 to 2012. Like biological variables, fishing intensity was averaged per sampling station.

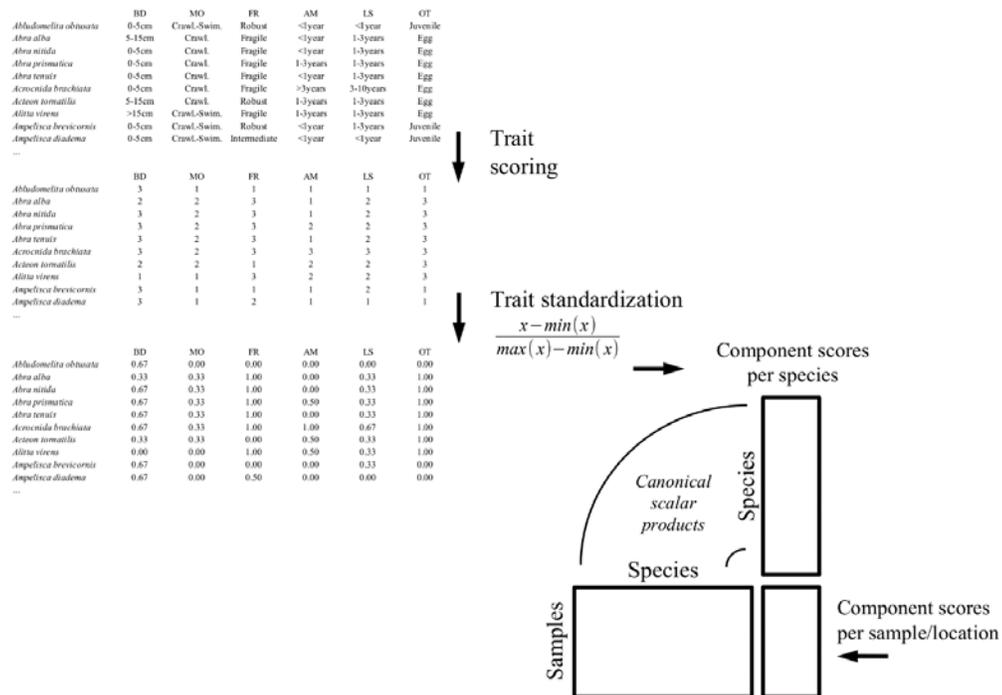


Figure 5.3.3. Computational procedure for the benthic community physical disturbance indicator. Modalities of initial qualitative trait data (upper matrix) are ascribed quantitative scores according to the degree of sensitivity to physical disturbance (see Table 1). After trait standardization and component computation per species, values are aggregated (component scores multiplied by field density and summed over the whole species community) in order to get synthetic scores per sample of field data.

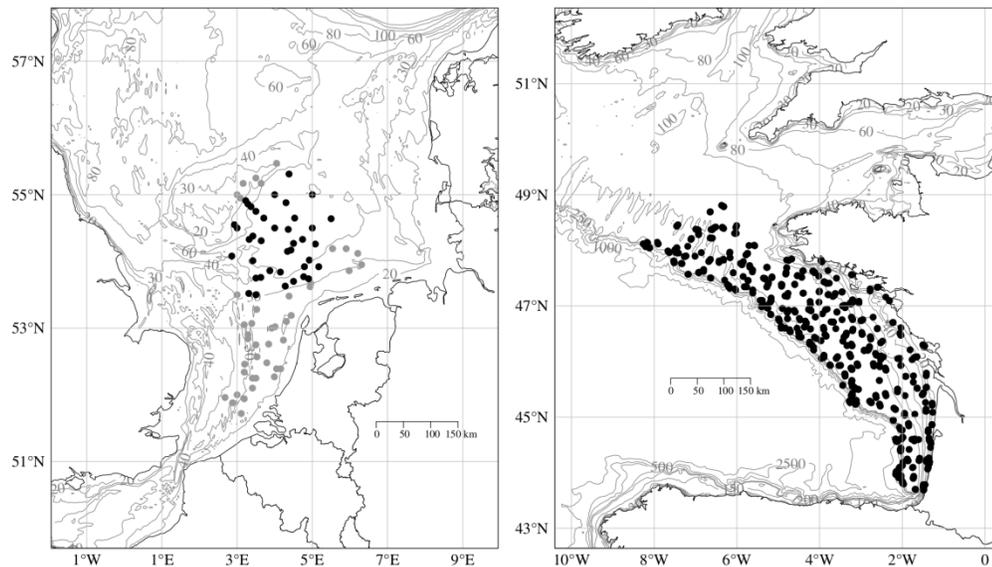


Figure 5.3.4. Case studies. Left - Dutch EEZ: grey dots, shallow sandy habitat; black dots, deep muddy habitat. Right - Bay of Biscay: the samples (black dots) within the French EEZ are included in the case study. Isobaths in both case studies are displayed by light grey lines.

5.3.5.2 Bay of Biscay

This case study analysed data on benthic invertebrate megafauna from the French Groundfish Survey in the Celtic Sea and Bay of Biscay (EVHOE). The survey has been conducted annually in autumn since 1992. The survey area ranges between 48°30'N in the North and the northern margin of Gouf de Cap Breton in the South (43°08'N; Figure 5.3.4). A 36/47 GOV trawl is used with a 20 mm mesh codend liner. On average 70 stations are sampled annually by GOV trawl. Non-commercial invertebrate species have been systematically recorded since 2010. The field data include species biomass and densities. However, since the two metrics are strongly correlated, only the densities were used for the case study. Species traits information come from the BENTHIS database (EU-project). As mentioned in section 5.3, some traits required by the indicator were missing from the database and were thus inferred. The computation of the benthic indicator was done using the procedure presented in Figure 5.3.3. For the purpose of this case study, 36 megabenthic species were included in the analyses representing respectively, 63 % and 69 % of the total density and biomass in the survey.

Fishing effort has been computed from the methodology for VMS data developed in the BENTHIS project (Eigaard *et al.*, 2017). Given that bottom trawl can affect the benthic fauna at various scales (cf. indicator components), the dataset of the fishing effort was integrated at several temporal and spatial scales. Fishing effort was defined as the sum or the average of the swept area (km²) observed in a pixel in a day, a month, two months, three months, six months and a year before the date of the scientific survey. Concurrently, a spatial integration was also done. It defined an area of influence around the sampling station of varying sizes, from the pixel in which the sampling occurred (i.e. 0 km) to a distance of 1, 2, or 5 km. The three components of the benthic indicator was tested against a combination of varying spatial and temporal scales of fishing effort integration.

5.3.5.3 Combining relevant traits information from various sources

Two trait data bases were used in the computations of the benthic indicator: one from ICES (2017) (Dutch waters) and a second one from the BENTHIS project for the Bay of Biscay. Information on longevity, burrowing depth and motility is common to both data bases, although modalities had to be matched and/or adjusted according to the case study in the Dutch EEZ (left hand side of Figure 5.3.5).

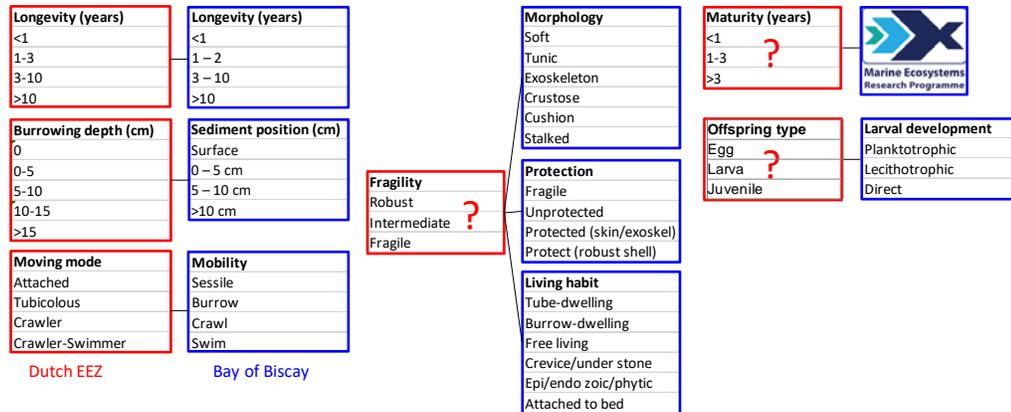


Figure 5.3.5. Combining relevant traits information from various sources.

Information on age at maturity, fragility and offspring type is not available for the Bay of Biscay study. Missing traits were therefore inferred. Age at maturity was obtained from the MERP Traits Explorer (). Fragility was inferred from trait information on morphology, protection and living habit, and offspring type was inferred from modalities of larval development (right hand side of Figure 5.3.3).

5.3.5.4 Distribution of indicator values

Figure 5.3.6 illustrates the distribution of the indicator components. In both case studies, distributions are skewed to the left due to the dominance of organisms with low sensitivity and vulnerability. Multi-modality of recovery components is apparent for the Dutch fauna with isolated small modes towards larger indicator values.

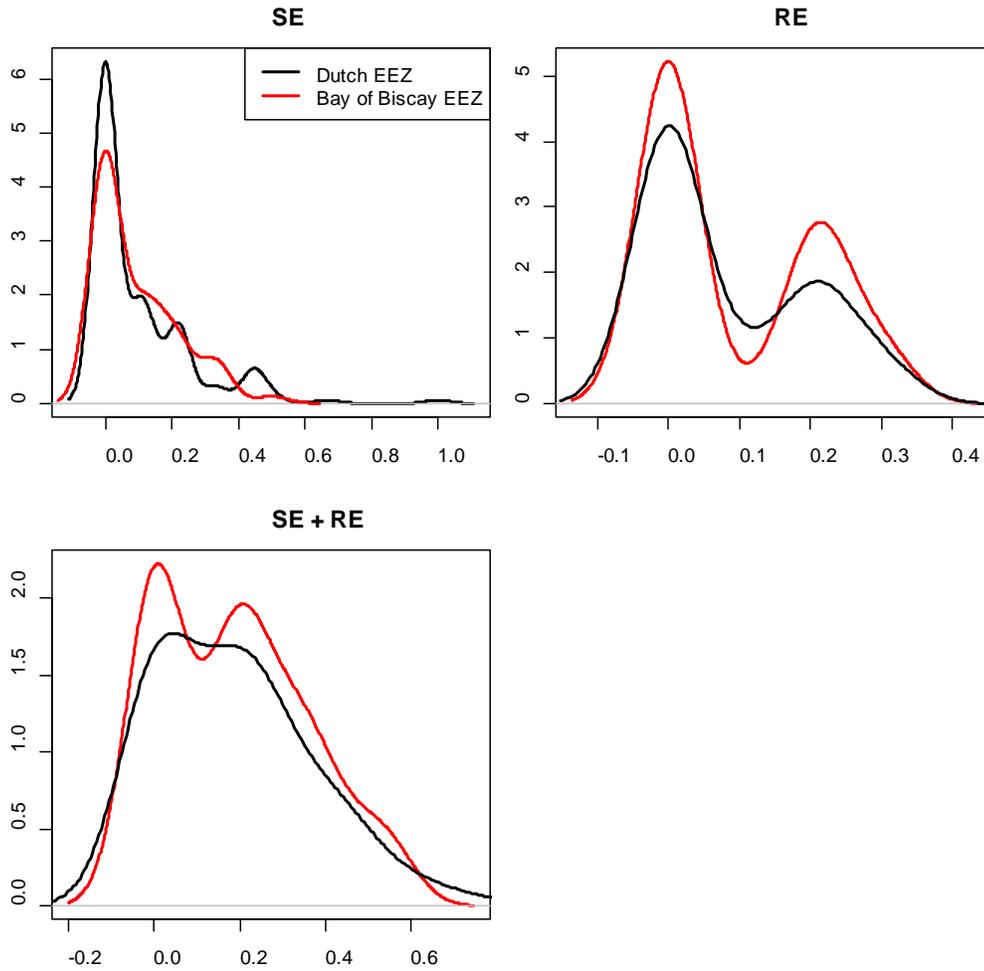


Figure 5.3.6. Comparative distributions of the benthic indicator components from the two case studies. The two upper graphs display the distribution of the sensitivity and recovery components. The bottom graph illustrates the distribution of the sum of both components. Distributions are displayed as density curves, i.e. a smooth approximation of the indices histograms.

5.3.5.5 Response of the benthic indicator to bottom trawling

a. Dutch EEZ

The outcomes from the application of the benthic indicator to the Dutch EEZ are consistent with theoretical predictions. Firstly, when removing confounding environmental effects, bottom trawling influenced the indicator components significantly (Table 5.3.3, Table 5.3.4 and Table 5.3.5; column “F|E”). Secondly, every significant response displayed in Figures 5.3.7 and 5.3.8 was negative, supporting the assumptions the indicator is based on.

Table 5.3.3. Variation partitioning of the different component combinations in the Dutch EEZ (whole area). Values represent the adjusted R² explaining the amount of variance of the response variable (SE, sensitivity component; RE, recovery component). Explanatory variables: E, environmental variables; F, fishing intensity; All, both; E|F, environment when removing fishing effect; E∩F, confounding effect; F|E, fishing when removing environmental effect. "ns", not significant. Note that the confounding effect E∩F cannot be tested due to its null degree of freedom (Legendre and Legendre, 2012).

Density	Response	E	F	All	E F	E∩F	F E
Absolute individual density	SE	0.708	0.255	0.715	0.461	0.248	ns
	RE	0.848	0.461	0.853	0.391	0.457	ns
	SE + RE	0.812	0.444	0.821	0.377	0.435	0.009
	SE × RE	0.876	0.476	0.878	0.402	0.474	ns
	SE + SE × RE	0.750	0.346	0.761	0.414	0.336	1.000
	RE + SE × RE	0.851	0.464	0.855	0.391	0.460	ns
	SE + RE + SE × RE	0.816	0.446	0.824	0.378	0.438	0.009
Absolute biomass density	SE	0.150	0.122	0.254	0.132	0.018	0.104
	RE	0.522	0.281	0.577	0.296	0.227	0.054
	SE + RE	0.464	0.271	0.532	0.261	0.202	0.069
	SE × RE	0.507	0.573	0.617	0.044	0.463	0.110
	SE + SE × RE	0.226	0.320	0.360	ns	0.186	0.134
	RE + SE × RE	0.511	0.303	0.570	0.267	0.243	0.060
	SE + RE + SE × RE	0.513	0.290	0.571	0.281	0.231	0.077
Absolute species richness	SE	0.827	0.346	0.824	0.478	0.348	ns
	RE	0.902	0.443	0.901	0.458	0.444	ns
	SE + RE	0.896	0.440	0.895	0.455	0.441	ns
	SE × RE	0.926	0.592	0.928	0.336	0.590	ns
	SE + SE × RE	0.865	0.412	0.863	0.452	0.414	ns
	RE + SE × RE	0.908	0.457	0.908	0.451	0.458	ns
	SE + RE + SE × RE	0.898	0.439	0.897	0.458	0.440	ns
Relative individual density	SE	0.364	ns	0.401	0.371	NA	0.037
	RE	0.829	0.394	0.836	0.443	0.386	0.008
	SE + RE	0.789	0.407	0.812	0.406	0.383	0.023
	SE × RE	0.829	0.477	0.835	0.358	0.472	ns
	SE + SE × RE	0.543	0.278	0.598	0.320	0.223	0.055
	RE + SE × RE	0.831	0.405	0.838	0.433	0.398	0.007
	SE + RE + SE × RE	0.795	0.415	0.817	0.401	0.394	0.021
Relative biomass density	SE	0.297	0.058	0.369	0.311	NA	0.073
	RE	0.577	0.445	0.618	0.173	0.404	0.041
	SE + RE	0.466	0.476	0.539	0.063	0.403	0.073
	SE × RE	0.624	0.608	0.662	0.054	0.570	0.038
	SE + SE × RE	0.294	0.273	0.383	0.110	0.183	0.089
	RE + SE × RE	0.587	0.475	0.629	0.154	0.433	0.042
	SE + RE + SE × RE	0.487	0.499	0.557	0.057	0.430	0.070
Relative species richness	SE	0.384	ns	0.382	0.359	0.025	ns
	RE	0.536	0.201	0.530	0.330	0.207	ns
	SE + RE	0.590	0.195	0.586	0.390	0.199	ns
	SE × RE	0.645	0.271	0.641	0.370	0.275	ns
	SE + SE × RE	0.530	0.118	0.529	0.411	0.119	ns
	RE + SE × RE	0.553	0.214	0.547	0.333	0.220	ns
	SE + RE + SE × RE	0.598	0.208	0.594	0.385	0.212	ns

Table 5.3.4. Variation partitioning of the different component combinations in the shallow sandy habitat from the Dutch EEZ. See legend of Table 5.3.3 for details.

Density	Response	E	F	All	E F	E∩F	FE
Absolute individual density	SE	0.704	0.189	0.695	0.506	0.198	ns
	RE	0.680	0.302	0.682	0.380	0.300	ns
	SE + RE	0.704	0.251	0.700	0.449	0.255	ns
	SE × RE	0.826	0.365	0.829	0.464	0.361	ns
	SE + SE × RE	0.720	0.210	0.712	0.502	0.218	ns
	RE + SE × RE	0.693	0.311	0.695	0.385	0.308	ns
	SE + RE + SE × RE	0.710	0.259	0.707	0.448	0.262	ns
Absolute biomass density	SE	0.455	ns	0.438	0.448	0.007	ns
	RE	0.339	ns	0.329	0.297	0.042	ns
	SE + RE	0.391	ns	0.380	0.361	0.030	ns
	SE × RE	0.496	0.269	0.480	0.211	0.284	ns
	SE + SE × RE	0.443	ns	0.426	0.402	0.041	ns
	RE + SE × RE	0.334	ns	0.324	0.283	0.051	ns
	SE + RE + SE × RE	0.446	ns	0.438	0.400	0.046	ns
Absolute species richness	SE	0.838	0.436	0.835	0.399	0.439	ns
	RE	0.813	0.528	0.807	0.279	0.534	ns
	SE + RE	0.833	0.519	0.827	0.309	0.524	ns
	SE × RE	0.956	0.680	0.955	0.275	0.681	ns
	SE + SE × RE	0.868	0.500	0.866	0.366	0.502	ns
	RE + SE × RE	0.835	0.550	0.830	0.280	0.555	ns
	SE + RE + SE × RE	0.831	0.516	0.826	0.310	0.521	ns
Relative individual density	SE	0.441	ns	0.521	0.492	0.000	0.079
	RE	0.347	0.150	0.329	0.179	0.168	ns
	SE + RE	0.402	0.135	0.434	0.299	0.103	ns
	SE × RE	0.800	0.461	0.795	0.334	0.466	ns
	SE + SE × RE	0.473	0.079	0.539	0.460	0.012	0.067
	RE + SE × RE	0.373	0.190	0.356	ns	0.208	ns
	SE + RE + SE × RE	0.414	0.169	0.441	0.272	0.142	ns
Relative biomass density	SE	ns	ns	ns	ns	0.000	ns
	RE	0.283	0.155	0.282	ns	0.156	ns
	SE + RE	0.268	0.164	0.259	ns	0.172	ns
	SE × RE	0.582	0.390	0.581	0.191	0.391	ns
	SE + SE × RE	ns	0.103	ns	ns	0.132	ns
	RE + SE × RE	0.341	0.195	0.342	ns	0.193	ns
	SE + RE + SE × RE	0.318	0.199	0.312	ns	0.205	ns
Relative species richness	SE	ns	ns	ns	ns	0.039	ns
	RE	0.337	0.163	0.317	ns	0.183	ns
	SE + RE	0.354	0.174	0.334	ns	0.195	ns
	SE × RE	0.683	0.292	0.682	0.390	0.293	ns
	SE + SE × RE	0.431	0.143	0.416	0.273	0.158	ns
	RE + SE × RE	0.371	0.183	0.352	ns	0.203	ns
	SE + RE + SE × RE	0.388	0.192	0.369	0.176	0.212	ns

Table 5.3.5. Variation partitioning of the different component combinations in the deep muddy habitat from the Dutch EEZ. See legend of Table 5.3.3 for details.

Density	Response	E	F	All	E F	E∩F	F E
Absolute individual density	SE	0.596	0.159	0.694	0.535	0.061	0.098
	RE	0.630	0.356	0.647	0.291	0.339	ns
	SE + RE	0.584	0.380	0.618	0.238	0.346	ns
	SE × RE	0.664	0.344	0.671	0.327	0.336	ns
	SE + SE × RE	0.514	0.334	0.595	0.262	0.252	0.082
	RE + SE × RE	0.634	0.356	0.650	0.294	0.340	ns
	SE + RE + SE × RE	0.590	0.379	0.621	0.243	0.347	ns
Absolute biomass density	SE	0.419	0.447	0.490	ns	0.376	0.071
	RE	ns	0.157	0.157	ns	0.000	0.157
	SE + RE	ns	0.196	0.196	ns	0.000	0.196
	SE × RE	0.478	0.568	0.564	ns	0.482	0.086
	SE + SE × RE	0.422	0.567	0.512	ns	0.477	0.090
	RE + SE × RE	ns	0.191	0.267	ns	0.118	ns
	SE + RE + SE × RE	ns	0.227	0.274	ns	0.151	0.076
Absolute species richness	SE	0.489	ns	0.473	0.438	0.051	ns
	RE	0.438	0.241	0.438	0.197	0.241	ns
	SE + RE	0.434	0.234	0.425	0.191	0.242	ns
	SE × RE	0.522	0.497	0.580	ns	0.439	0.058
	SE + SE × RE	0.543	0.148	0.529	0.381	0.162	ns
	RE + SE × RE	0.456	0.270	0.461	0.191	0.265	ns
	SE + RE + SE × RE	0.450	0.259	0.446	0.188	0.262	ns
Relative individual density	SE	0.704	ns	0.720	0.745	0.000	ns
	RE	0.727	0.255	0.720	0.465	0.262	ns
	SE + RE	0.697	0.292	0.693	0.400	0.296	ns
	SE × RE	0.731	0.265	0.724	0.459	0.272	ns
	SE + SE × RE	0.494	0.389	0.521	ns	0.363	ns
	RE + SE × RE	0.730	0.257	0.723	0.466	0.264	ns
	SE + RE + SE × RE	0.702	0.290	0.698	0.408	0.294	ns
Relative biomass density	SE	0.697	0.438	0.722	0.284	0.412	ns
	RE	0.286	0.354	0.349	ns	0.291	0.063
	SE + RE	0.317	0.423	0.386	ns	0.354	0.069
	SE × RE	0.604	0.584	0.642	ns	0.546	ns
	SE + SE × RE	0.641	0.563	0.678	0.115	0.526	0.037
	RE + SE × RE	0.324	0.390	0.387	ns	0.327	ns
	SE + RE + SE × RE	0.352	0.448	0.418	ns	0.381	ns
Relative species richness	SE	0.422	ns	0.403	0.346	0.076	ns
	RE	ns	ns	ns	ns	0.000	ns
	SE + RE	ns	ns	ns	ns	0.014	ns
	SE × RE	ns	ns	ns	ns	0.037	ns
	SE + SE × RE	0.361	ns	0.340	0.347	0.013	ns
	RE + SE × RE	ns	ns	ns	ns	0.000	ns
	SE + RE + SE × RE	ns	ns	ns	ns	0.009	ns

At both spatial scales, densities derived from biomass were the most responsive (Tables 5.3.3 and 5.3.5), although very few responses were significant within the shallow sandy habitat (Table 5.3.4). Conversely, almost no significant relationship was found for the responses derived from absolute and relative species richness (computation based on presence/absence matrix). Globally, the amount of explained variance by the conditional fishing effect (F|E) is low due to strong confounding effects of environment and fishing intensity (E∩F).

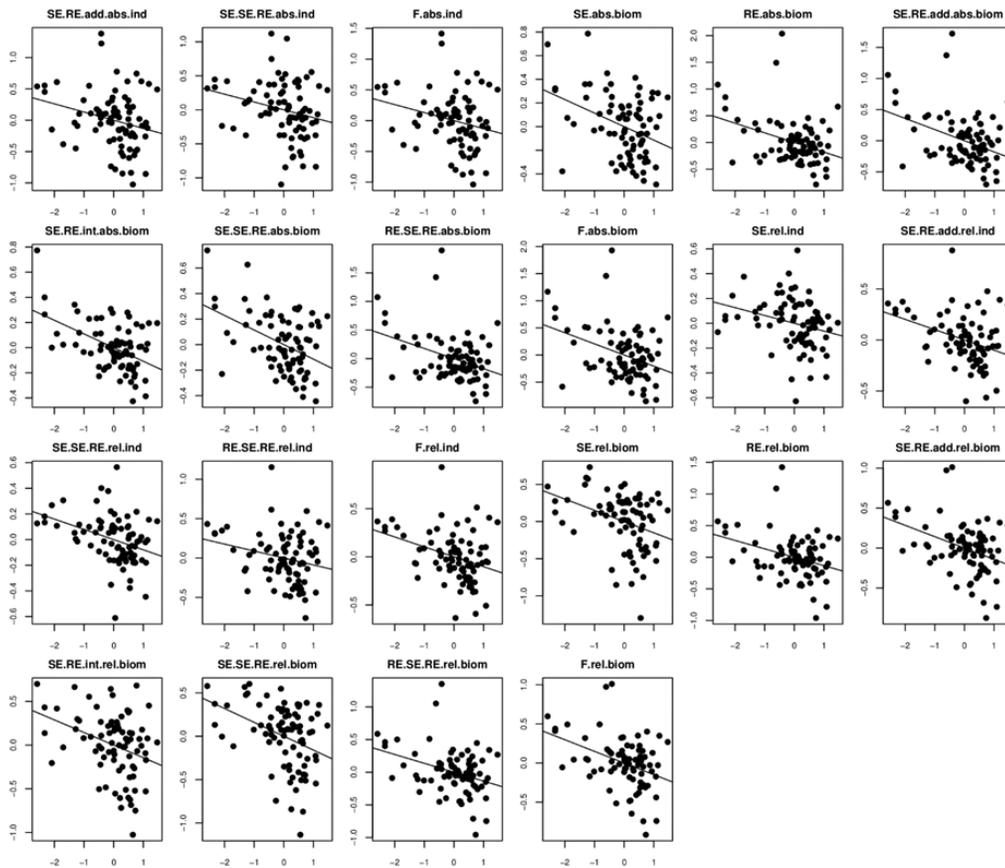


Figure 5.3.7. Significant relationships between the components of the benthic indicator and fishing intensity over the entire Dutch EEZ (see Table 5.3.3). Graph titles: SE, sensitivity component; RE, recovery component; “add” for additivity, “int” for interactivity and “F” for full formula in Tables 3, 4 and 5; “abs” for absolute faunal density, “rel” for relative faunal density; “ind” and “biom” for respectively individual and biomass density.

The responsiveness of the two individual components SE and RE are relatively even when considering the overall results, neither of the two being more often significant than the other. A few significant interactive effects were found, but they do not seem to increase the amount of explained variance, individually considered or added to other components. This suggests a relative functional independence of both sensitivity and recovery components. For instance, the highest amount of variance explained by F|E was found for the additive effect of both components in the deep muddy habitat ($R^2 = 0.20$, Table 5.3.5). Although a significant interactive effect was found ($R^2 = 0.09$), the full formula yields a lower amount of explained variance ($R^2 = 0.08$).

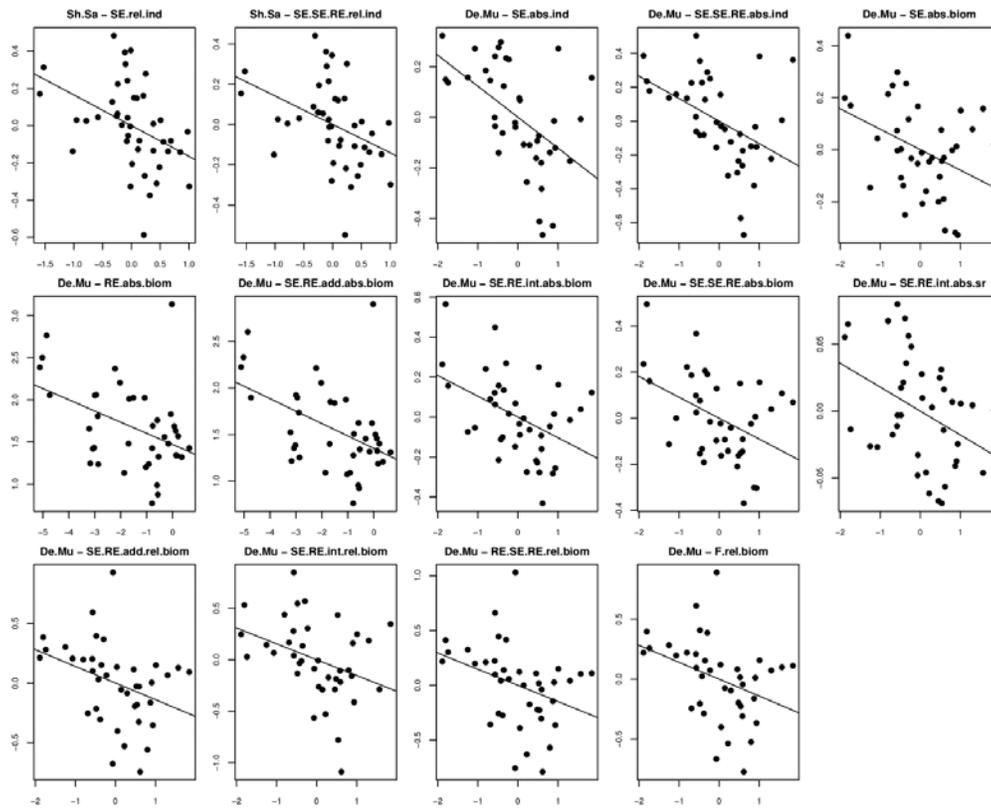


Figure 5.3.8. Significant relationships between components of the benthic indicator and fishing intensity within shallow sandy (“Sh.Sa”) and deep muddy (“De.Mu”) habitats of the Dutch EEZ (see Tables 5.3.4 and 5.3.5). See Figure 5.3.5 for description of remaining abbreviations.

b. Bay of Biscay

The outcomes from the application of the benthic disturbance to the French EEZ in the Bay of Biscay are also consistent with the theoretical predictions. There was no significant relationship between the benthic indicator and the bathymetry (all p -values > 0.05). Thus, the relationship between the benthic indicator and the fishing effort (at multiple spatial and temporal scales) was done directly, without partitioning the variance with the environment. Given that multiple tests were conducted, we corrected the α threshold to account for Type-2-error (i.e. incorrectly retaining a false result). The effect of bottom trawling on the indicator components was significant at different scales (Table 5.3.6). The low amount of variance explained may be due to the relative low number of species included in the analyses and notably the absence of species mostly sensitive to trawling (e.g. sea pens). Nevertheless, all the significant responses were as expected, systematically decreasing with fishing effort (Figure 5.3.9).

The RE component responded more strongly to the fishing effort than the SE component in most of the significant relationships, suggesting that the benthic community included in the case study has been already exposed to chronic trawling pressure.

Table 5.3.6. Amount of variability (R²) explained by fishing effort at varying temporal (0 to 12 months: 0 M, 3 M, 6 M and 12 M) and spatial (1 km and 5 km) integration (sum or mean) for the different component combinations of the benthic indicator in the French EEZ.

	SE	RE	SE + RE	SE * RE
VMS1km6Msum	ns	0.12	0.06	ns
VMS1km6Mmean	ns	0.11	0.03	ns
VMS5km0Msum	0.01	0.08	0.04	ns
VMS5km0Mmean	0.01	0.05	0.01	ns
VMS5km3Msum	ns	ns	ns	0.02
VMS5km6Msum	ns	ns	ns	0.04
VMS5km6Mmean	ns	ns	ns	0.06
VMS5km12Msum	ns	0.10	ns	0.05
VMS5km12Mmean	ns	0.12	ns	0.05

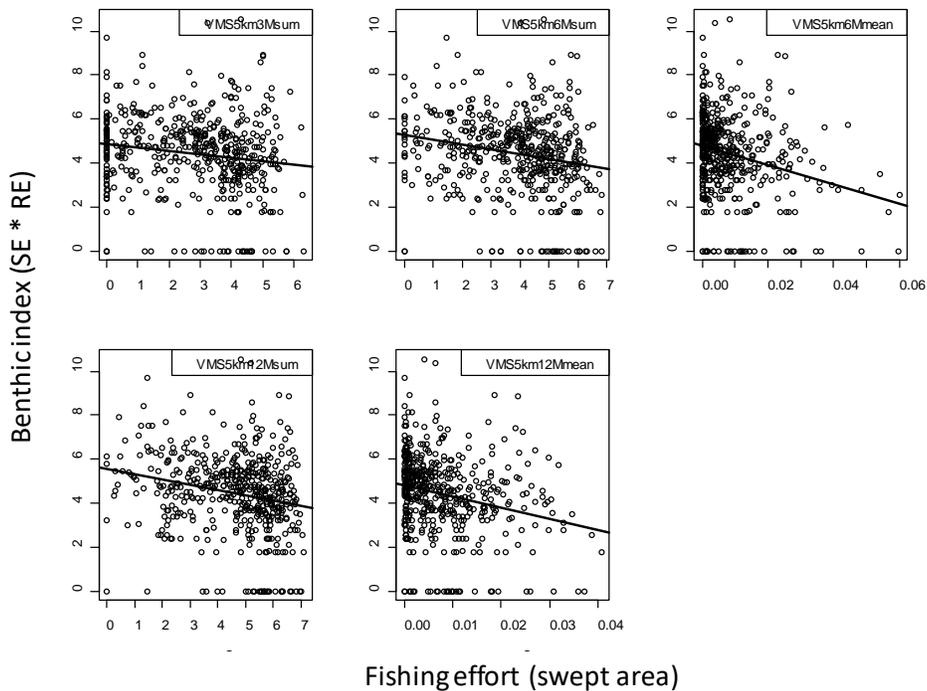


Figure 5.3.9. Examples of relationship between the fishing effort (at various spatio-temporal scales) and the benthic indicator in the French EEZ of the Bay of Biscay.

Projection of the benthic indicator was done for the French EEZ in the Bay of Biscay using the interpolated maps of species developed by Ifremer (Ifremer, 2015). In this spatial representation the density of each species is interpolated over the entire sampling area of the scientific survey. For each species, the score of the benthic indicator is weighted by its density in each pixel and the sum of all the weighted scores is projected in a benthic sensitivity map (Figure 5.3.10). For the time being, one has to keep in mind that the map is based on only 36 species and that although general patterns can be observed, careful interpretation must be done given the need to add supplementary sensitive species.

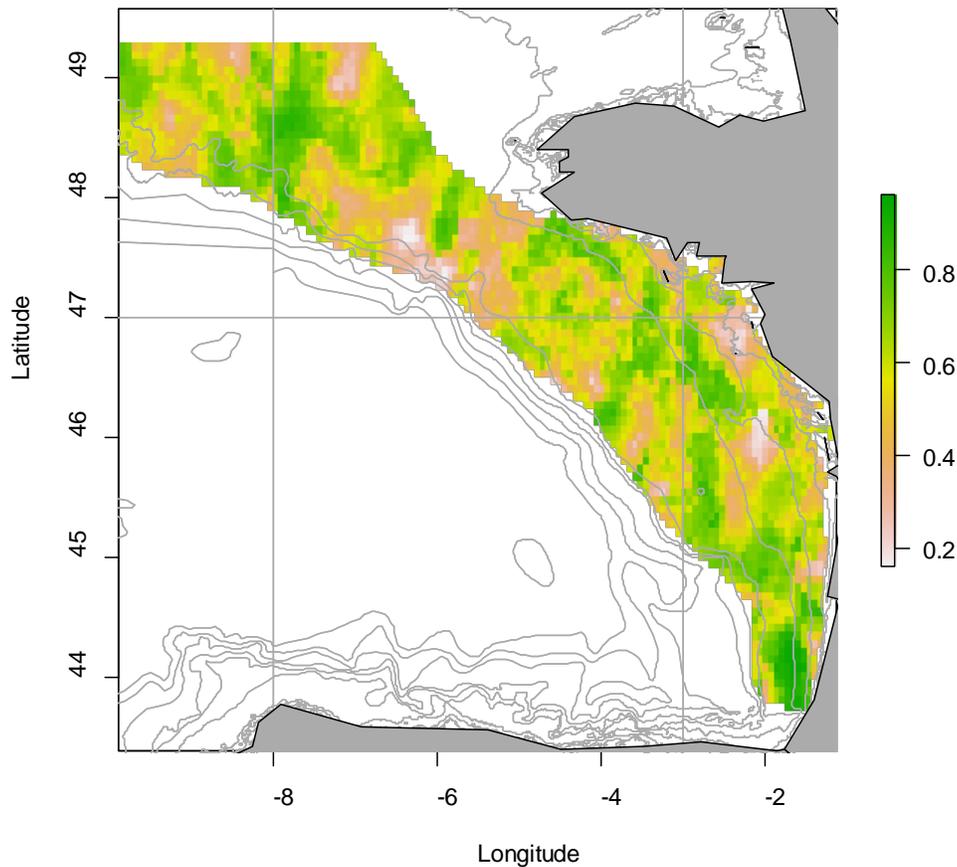


Figure 5.3.10. Spatial distribution of the Benthic indicator in the Bay of Biscay based on the analysis of the 36 megabenthic species. The map is the sum of each species interpolated density distribution weighted by its benthic indicator. The indicator is proportional to the sensitivity to trawling, i.e. the higher the indicator values, the more sensitive the area concerned.

5.3.6 Discussion

Distribution of indicator values

Our case study data sets are dominated by species that are relatively resilient or resistant to physical disturbance (i.e. low vulnerability) whereas there are only few vulnerable

species. This suggests that achieving a high indicator score would be possible in habitats dominated by these few vulnerable species. We would expect that applying the indicator in broader geographical/spatial/temporal contexts to reveal more diverse distributions of indicator values. It is possible that the skewed distribution of indicator values in our case studies is partly a result of several decades of intensive bottom trawling, and that another distributional pattern would have been expected under pristine conditions.

Indicator sensitivity

The distribution of indicator values (see above) suggests that even in habitats that have been exposed to chronic trawling pressure, the two components are sufficiently sensitive to detect differences in the effects of varying fishing intensities on benthic communities. The two habitats within the Dutch EEZ offer an interesting experimental design since resident species communities are differently affected by bottom trawling intensity (ICES, 2017). The results obtained for the SE and RE components corroborate this observation. Highest significant responses were observed in the deep muddy habitat where communities are more vulnerable although exposed to lower trawling intensities (ICES, 2017). The shallow sandy habitat is characterized by natural stress imposed by high water velocities and mobile sands that favour species that are resistant and resilient to frequent natural disturbance. The two sole significant responses from this habitat, SE and SE+SE×RE (based on relative individual density), reflect the resilience of the fauna (Table 5.3.2).

The lack of significant responses from species richness suggests that all combination of SE and RE values can be encountered in any community. However, species rich communities are generally encountered under environmental conditions that benefit the settlement of most species. Biotic interactions and rare disturbance events are then the primary factors controlling the number of species. The Dutch deep muddy habitat is a typical example of such communities. Here, most species mature late, so that we would expect RE to respond significantly. In general, most organisms in species-rich benthic communities tend to occur close to the sediment surface, with possibly high SE values. The long-term trawling pressure on the habitats studied could simply have selected the less exposed species so that SE and RE functionalities could remain similar over the whole species richness gradient.

Experimental constraints

The case study areas have been affected by bottom trawling for several decades, which imposes strong experimental constraints on testing the indicator. An additional constraint is the presence of large confounding effects that precludes reliable estimates of the real amount of variance explained by fishing. This results in low variances purely explained by fishing intensity because the fishing effort follows the natural gradients (i.e. primary productivity). However, this might not apply elsewhere. Our result does not imply that the benthic indicator performs badly, as it is sufficiently sensitive to detect differences in trawling pressure against a background of environmental variability (see above).

Removal of environmental variance

The slope of the relationship between indicator values and fishing pressure is not affected by the removal of environmental variance (in the computational procedure of varia-

tion partitioning from Legendre and Legendre (2012)). The explained variance in the variation partitioning accounts for the likelihood of trawling effect, but is constrained by confounding effect, whereas the slope of the relationship might be more pertinent to assess the responsiveness of the indicator components. Future investigation should focus on this aspect.

5.3.7 The future

Future investigations of WGBIODIV will focus on comparison of indicator values of benthic invertebrate assemblages (infauna derived from grabs and cores, and megafauna derived from trawls) from spatially extensive surveys. WGBIODIV will carry out meta-analyses of these data to compare the distribution of indicator values between different geographical areas and, for some locations, over time. We intuitively expect that assemblages from heavily trawled areas will have higher proportions of organisms resilient to trawling disturbance than areas where trawling pressure is lower, but we have no *a priori* expectation of the precise nature of trawl effects (i.e. the distribution of indicator values).

Specific hypotheses that could be tested include:

H₁: Assemblages of the same broad group of animals (i.e. benthic infauna or benthic megafauna) do not respond differently to trawl disturbance (i.e. there is a generalised response with the proportion of resilient organisms increasing with increasing trawling pressure).

H₂: The indicator distribution differs according to the environmental conditions that the assemblage normally experiences (i.e. high proportion of resilient organisms in areas that experience high levels of natural physical disturbance).

H₃: Assemblages from different habitats, experiencing different levels of natural physical disturbance, do not respond differently to trawl disturbance (i.e. there is a generalised response with the proportion of resilient organisms increasing with increasing trawling pressure).

5.3.8 References

- Balmford, A., Bruner, A., Cooper, P., Costanza, R., Farber, S., Green, R.E., Jenkins, M., Jerfferiss, P., Jessamy, V., Madden, J., Munro, K., Myers, N., Naeem, S., Paavola, J., Rayment, M., Rosendo, S., Roughgarden, J., Trumper, K. and Turner, R.K. 2002. Economic reasons for conserving wild nature. *Science*, 297: 950-953.
- Beauchard O., Verissimo H., Queirós A.M., Herman P.M.J., 2017. The use of multiple biological traits in marine community ecology and its potential in ecological indicator development. *Ecological indicators* 76:81–96.
- Dencker, T.S., Pecuchet, L., Beukhof, E., Richardson, K., Payne, M.R., Lindegren, M. 2017. Temporal and spatial differences between taxonomic and trait biodiversity in a large marine ecosystem: Causes and consequences. *PLoS ONE* 12(12): e0189731.
- Eigaard, O.R., Bastardie, F., Hintzen, N.T., Buhl-Mortensen, L., Buhl-Mortensen, P., Catarino, R., Dinesen, G.E., Egekvist, J., Fock, H.O., Geitner, K. and Gerritsen, H.D. 2016. The footprint of bottom trawling in European waters: distribution, intensity, and seabed integrity. *ICES Journal of Marine Science*, 74(3): 847–865.

- Gregory, A.J., Atkins, J.P., Burdon, D. Elliott, M.. 2013. A problem structuring method for ecosystem-based management: The DPSIR modelling process. *European Journal of Operational Research* 227(3): 558–569.
- Halpern, B.S., Walbridge, S., Selkoe, K.A., Kappel, C.V., Micheli, F., D'Agrosa, C., Bruno, J.F., Casey, K.S., Ebert, C., Fox, H.E., Fujita, R., Heinemann, D., Lenihan, H.S., Madin, E.M.P., Perry, M.T., Selig, E.R., Spalding, M., Steneck, R., and Watson, R. 2008. A global map of human impact on marine ecosystems. *Science* 319: 948-952.
- ICES 2015. Report of the ICES Working Group on Biodiversity Science (WGBIODIV). ICES CM 2015/SSGEPD:04, 310 pp.
- ICES, 2017. Report of the Workshop to evaluate regional benthic pressure and impact indicator(s) from bottom fishing (WKBENTH), 28 February–3 March 2017, Copenhagen, Denmark. ICES CM 2017/ACOM:40. 233 pp.
- Ifremer, 2015. Atlas des invertébrés Benthiques issus des campagnes de surveillance halieutique de l'Ifremer. <http://atlasbenthal.ifremer.fr>
- Jennings, S. 2004. The ecosystem approach to fishery management: a significant step towards sustainable use of the marine environment? *Marine Ecology Progress Series* 274: 279-282.
- Jørgensen, C.K., Enberg, E.S., Dunlop, R., Arlinghaus, D.S., Boukal, K., Brander, B., Ernande *et al.* 2007. Managing the world's evolving fish stocks. *Science* 318: 1247–1248.
- Kaiser, M.J., Clarke, K.R., Hinz, H., Austen, M.C.V., Somerfield, P.J. & Karakassis, I. 2006. Global analysis of response and recovery of benthic biota to fishing. *Marine Ecology Progress Series* 311: 1- 14.
- Kaiser, M.J., Ramsay, K., Richardson, C.A., Spence, F.E., Brand, A.R. 2000. Chronic fishing disturbance has changed shelf sea benthic community structure. *Journal of Animal Ecology* 69: 494-503.
- Laurila-Pant, M., Lehikoinen, A., Uusitalo, L. Venesjärvi, R. 2015. How to value biodiversity in environmental management?. *Ecological indicators* 55: 1–11.
- Lavorel S., Grigulis K., McIntyre S., Williams N.S.G., Garden D., Dorrough J., Berman S., Quétier F., Thébault A., Bonis A., 2008. Assessing functional diversity in the field – methodology matters! *Functional Ecology* 22:134–147.
- Legendre P., Legendre L., 2012. Numerical ecology, Third Edition. Elsevier, 1006 p.
- Pitcher C.R., Ellis N., Jennings S., Hiddink J.G., Mazor T., Kaiser M.J., Kangas M.I., McConnaughey R.A., Parma A.M., Rijnsdorp A.D., Suuronen P., Collie J.S., Amoroso R., Hughes K.M., Hilborn R., 2016. Estimating the sustainability of towed fishing-gear impacts on seabed habitats: a simple quantitative risk assessment method applicable to data-limited fisheries. *Methods in Ecology and Evolution* 8:472–480.
- Watson, R., Revenga, C., Kura, Y. 2006. Fishing gear associated with global marine catches: II. Trends in trawling and dredging. *Fisheries Research* 79(1–2): 103–111.
- Worm, B., Barbier, E.B., Beaumont, N., Duffy, J.E., Folke, C., Halpern, B.S., Jackson, J.B.C., Lotze, H.K., Micheli, F., Palumbi, S.R., Sala, E., Selkoe, K.A., Stachowicz, J.J., and Watson, R. 2006. Impacts of biodiversity loss on ocean ecosystem services. *Science* 314: 787-790.

6 Cooperation

During the 2016–2018 working cycle, WGBIODIV was readjusting its objectives. There has been no outreach to other ICES working groups, which could be improved once more products (indicators, maps, study results) from the current and future ToRs (2019–2021) are available.

7 Summary of Working Group self-evaluation and conclusions

WGBIODIV decided to issue a resolution for another three-year working period to complete the promising work on the benthic response indicators, diversity of trophic guilds and indicator development protocol.

Annex 1: Lists of participants

WGBIODIV 08–12.02.2016, Donostia (San Sebastian), Spain

Name	Country	Country	E-mail
Anaïs Aubert		France	anais.aubert@mio.osupytheas.fr
Andrea Belgrano		Sweden	andrea.belgrano@slu.se
Ángel Borja		Spain	aborja@azti.es
Christopher Lynam		UK	chris.lynam@cefas.co.uk
Felipe Artigas		France	felipe.artigas@univ-littoral.fr
GerJan Piet		Netherlands	gerjan.piet@wur.nl
Henrike Rambo		Germany	henrike.rambo@ti.bund.de
Isabelle Rombouts		France	isabelle.rombouts@univ-lille1.fr
Maite Louzao		Spain	mlouzao@azti.es
Michaela Schratzberger		UK	michaela.schratzberger@cefas.co.uk
Nikolaus Probst		Germany	nikolaus.probst@ti.bund.de
Olivier Beauchard		Netherlands	olivier.beauchard@nioz.nl
Oscar Bos		Netherlands	oscar.bos@wur.nl
Simon Greenstreet		UK	s.greenstreet@marlab.ac.uk
Tim Spaanheden Dencker		Denmark	tspa@aqua.dtu.dk



ICES WGBIODIV 2016 (from left to right): Simon Greenstreet, Henrike Seidel, Nikolaus Probst, Felipe Artigas, Anaïs Aubert, Ángel Borja, Tim Spaanheden Dencker, Michaela Schratzberger, Gerjan Piet, Christopher Lynam, Oliver Beauchard, Andrea Belgrano, Oscar Bos.

WGBIODIV, 06–10.02.2017, Venice, Italy

NAME	AFFILIATION	EMAIL
Andrea Belgrano	Swedish University of Agricultural Sciences (SE)	andrea.belgrano@slu.se
Anik Brind'Amour	Ifremer (F)	Anik.Brindamour@ifremer.fr
Christopher Lynam	Cefas (UK)	Chris.Lynam@cefass.co.uk
Fabio Pranovi	Ca' foscari University of Venice (I)	fpranovi@unive.it
Felipe Artigas	Université Littoral (F)	felipe.artigas@univ-littoral.fr
Gerjan Piet	Wageningen Marine Research (NL)	Gerjan.piet@wur.nl
Gert Van Hoey	ILVO (B)	Gert.Vanhoey@ilvo.vlaanderen.be
Henrik Gislason	DTU Aqua (DK)	hg@aquadtu.dk
Hilde Trannum	NIVA (N)	hilde.trannum@niva.no
Isabelle Rombouts	MNHN (F)	isabelle.rombouts@univ-lille1.fr
Laura Pettit	JNCC (UK)	Laura.Pettit@jncc.gov.uk
Laura Robson	JNCC (UK)	Laura.Robson@jncc.gov.uk
Michaela Schratzberger	Cefas (UK)	michaela.schratzberger@cefass.co.uk
Nik Probst	Von Thünen Institute (D)	nikolaus.probst@thuenen.de
Olivier Beauchard	VLIZ (B)	Olivier.beauchard@vliz.be
Oscar Bos	Wageningen Marine Research (NL)	Oscar.bos@wur.nl
Paul Somerfield	Plymouth Marine Lab (UK)	pjso@pml.ac.uk
Saša Raicevich	Institute for Environmental Protection and Research, Rome (I)	sasa.raicevich@isprambiente.it
Simon Greenstreet	Marine Lab Schotland (UK)	s.greenstreet@marlab.ac.uk



ICES WGBIODIV 2017 (from left to right): Oscar Bos, Henrik Gislason, Christopher Lynam, Simon Greenstreet, Isabelle Rombouts, Andrea Belgrano, Laura Robson, Laura Pettit, Gerjan Piet, Olivier Beauchard, Hilde Trannum, Anik Brind'Amour, Felipe Artigas, Paul Somerfield, Gert van Hoey, Nikolaus Probst (not on this photo: Saša Raicevich and Fabio Pranovi).

WGBIODIV, 05–09.02.2018, ICES Headquarter, Copenhagen, Denmark

Name	Institute (country)	Email
Andrea Belgrano	Swedish University of Agricultural Sciences (SE) and Swedish Institute for the Marine Environment (SIME) (SE)	andrea.belgrano@slu.se
Anik Brind'Amour	Ifremer (F)	Anik.Brindamour@ifremer.fr
Christopher Lynam	Cefas (UK)	Chris.Lynam@cefas.co.uk
Felipe Artigas	Université Littoral (F)	felipe.artigas@univ-littoral.fr
Henrike Rambo	Von Thünen Institute (D)	Henrike.rambo@thuenen.de
Michaela Schratzberger	Cefas (UK)	michaela.schratzberger@cefas.co.uk
Nikolaus Probst	Von Thünen Institute (D)	nikolaus.probst@thuenen.de
Olivier Beauchard	VLIZ (B)	Olivier.beauchard@vliz.be
Oscar Bos	Wageningen Marine Research (NL)	Oscar.bos@wur.nl
Paul Somerfield	Plymouth Marine Lab (UK)	pjso@pml.ac.uk
Simon Greenstreet	Marine Lab Schotland (UK)	s.greenstreet@marlab.ac.uk
Victoria Magath	University of Hamburg (D)	victoria.magath@uni-hamburg.de



ICES WG BIODIV 2018 (from left to right): Oscar Bos, Andrea Belgrano, Felipe Artigas, Christopher Lynam, Nikolaus Probst, Vicky Magath, Paul Somerfield, Anik Brind'Amour, Michaela Schratzberger, Simon Greenstreet, Olivier Beauchard (not in this photo: Henrike Rambo (photo: Terhi Minkkinen, ICES)).

Annex 2: Recommendations

WGBIODIV does not issue any recommendations as a result from the work achieved between 2016 and 2018, as further work is needed to consolidate the current achievements.

Annex 3: WGBIODIV draft resolution 2019–2021

The **Working Group on Biodiversity Science** (WGBIODIV), chaired by Christopher Lynam, United Kingdom, and Andrea Belgrano, Sweden, will work on ToRs and generate deliverables as listed in the Table below.

	Meeting dates	Venue	Reporting details	Comments (change in Chair, etc.)
Year 2019	4–8 February	Copenhagen, Denmark	Interim report by 1 May	
Year 2020			Interim report by	
Year 2021			Final report by	

ToR descriptors

ToR	Description	Background	Science priorities	Duration	Expected Deliverables
a	<p>Test the benthic response indicator:</p> <p>Capacity to support formal assessment and management advice</p> <p>Use the indicator to explore its effectiveness in different disturbance/environmental change scenarios</p> <p>Explore the utility of the indicator in a broader geographical context</p>	<p>In the previous three-year term WGBIODIV concluded that indicators to assess biodiversity are not working because the lack a theoretical foundation. WGBIODIV addressed this problem by designing a trait-based sensitivity indicator of benthic communities. For example, establishment of pressure-state relationships, validation of indicator calculations and evaluation against the WGBIODIV indicator quality criteria is still pending. Indicator testing is the logical next step following the successful completion of the previous ToR.</p>		3 years	A tested and operational indicator of community response to ecosystem change.
b	<p>Investigate mechanisms linking trophic guilds under contrasting levels of pressure and/or primary production in case study areas:</p> <p>Using diet/trait information, and both predator and prey abundance to estimate potential impact on prey due to consumption by predators.</p>	<p>Understanding of pressure-state relationships are fundamental to indicator assessments. However, as pressure is removed through management and ecosystems begin to recover, the nature and/or strength of previously defined pressure-state relationships may</p>		3 years	Identify whether recovery of ecosystem components (e.g. predatory fish) can lead to depletion of prey groups such that natural processes dominate change.

	<p>Contrast risk due to natural mortality (consumption) with risk due to fishing pressure</p> <p>Project change in risk for prey groups due to increase in predator abundance or shifts in community composition as predator communities recover</p> <p>Clearly define roles of top down control and bottom up limitation at different trophic levels</p>	<p>change.</p> <p>Climate change effects may further modify or mask the effects of anthropogenic pressures.</p> <p>This ToR will investigate responsiveness of indicators to pressure in regional seas where demersal fishing pressure has been reduced and temperature has increased.</p>		
c	<p>Examine the efficacy of spatial management measures as means of conserving, protecting and promoting marine biodiversity</p>	<p>The implementation of the management plans for the Natura 2000-sites is under way and will have substantial impacts on human activities, namely by spatial measures such as (partial) fisheries closures and marine reserves.</p> <p>However, the Habitat Directive addresses only a limited range of taxa i.e. excluding the majority of epibenthic species and marine fish. WGBIODIV considers it as important to know, how much the current MPA networks will contribute to the protection of these taxa.</p>	3 years	<p>Maps of biodiversity in selected marine regions to inform on occurrence of biodiversity and to guide spatial management for its conservation.</p>
d	<p>Develop indicators on biological functioning</p>	<p>Requested by SCICOM in January 2018</p>	3 years	<p>WGBIODIV requests SCICOM to draft a precise ToR or to provide more detailed background, so that the group understands what is being asked of them.</p>

Summary of the Work Plan

Year 1	<p>Develop assessment targets for benthic response indicator; provide first analysis on trophic guilds and linkages to pressures; develop method to create and overlay single-species distributions</p>
Year 2	<p>Final evaluation of benthic response indicator; progress analysis of trophic guilds vs. anthropogenic pressures; create maps of biodiversity hotspots</p>
Year 3	<p>Finalise and evaluate work on trophic guild and hotspots</p>

Supporting information

Priority	The current activities of this group will lead ICES into issues related to the integrated ecosystem assessments and the implementation of the ecosystem approach to marine management. Consequently, these activities are considered to have a very high priority.
Resource requirements	The research programmes which provide the main input to this group are already underway, and resources are already committed. The additional resource lies in the commitment of scientist from different member states to participate in the group.
Participants	Participation in WGBIODIV has slightly increased due to the outreach strategy of hosting meeting in Spain and Italy, thereby attracting scientist from host countries and Mediterranean area.
Secretariat facilities	None.
Financial	No financial implications.
Linkages to ACOM and groups under ACOM	There is a linkage to ACOM/SCICOM steering group Integrated Ecosystem assessments (IEA). The results of WGBIODIV are important to WGECO and may be of relevance for WGINOSE and WGIAB.
Linkages to other committees or groups	The outcomes of WGBIODV will be important to the ICES high priority work area 'Marine Strategy Framework Directive (MSFD)'.
Linkages to other organizations	OSPAR, HELCOM, European Commission

Annex 4: WGBIODIV self-evaluation

- 1) Working Group name:
Working Group on Biodiversity Science
- 2) Year of appointment:
2015
- 3) Current Chairs:
Oscar Bos, Nikolaus Probst
- 4) Venues, dates and number of participants per meeting:
08–12.02.2016 San Sebastian (15),
06–10.02.2017, Venice (19),
05–09.02.2018, Copenhagen (12)

WG Evaluation

- 5) If applicable, please indicate the research priorities (and sub priorities) of the Science Plan to which the WG make a significant contribution.

1 - Assess the physical, chemical and biological state of regional seas and investigate the predominant climatic, hydrological and biological features and processes that characterise regional ecosystems

5 - Quantify the role of structural and functional diversity in marine ecosystems in providing stability and resilience

6 - Investigate linear and non-linear ecological responses to change, the impacts of these changes on ecosystem structure and function and their role in causing recruitment and stock variability, depletion and recovery

7 - Develop end to end modelling capability to fully integrate natural and anthropogenic forcing factors affecting ecosystem functioning

9 - Identify indicators of ecosystem state and function for use in the assessment and management of ecosystem goods and services

11 - Develop methods to quantify multiple direct and indirect impacts from fisheries as well as from mineral extraction, energy generation, aquaculture and other anthropogenic activities and estimate the vulnerability of ecosystems to such impacts.

14 - Evaluate ecological, economic and social trade-offs between ecosystem protection and sustainable use to advise on management of human activity in marine ecosystems

15 - Develop tactical and strategic models to support short and long term fisheries management and governance advice and increasingly incorporate spatial components in such models to allow for finer scale management of marine habitats and populations

16 - Quantify and map biological, ecological and environmental values with an aim to optimize ecosystem use and minimize environmental impacts in relation to ecosystem carrying capacity

18 - Identify objectives for IEA's that address ecosystem stability and health, taking cognizance of ecological, social and economic sustainability goals as well as multi scale issues.

19 - Identify issue based ecosystem questions relevant to science and management needs that can be addressed by developing IEA's

22 - Determine and demonstrate what modelling and analytical approaches will allow projections of ecosystem states in IEA's

6) In bullet form, list the main outcomes and achievements of the WG since their last evaluation. Outcomes including publications, advisory products, modelling outputs, methodological developments, etc.

- Development of benthic response indicator
- Review and analysis on impacts of sampling effort on diversity metrics
- Generic protocol on conceptual indicator development
- Two publications
 - *Probst, W. N., Lynam, C. P. 2016. Aggregated assessment results depend on aggregation method and framework structure - a case study within the European Marine Strategy Framework Directive. Ecological Indicators, 61: 871–881.*
 - *Rambo, H., Stelzenmueller, V., Greenstreet, S.P.R., Moellmann, C. 2017. Mapping fish community biodiversity for European marine policy requirements. ICES Journal of Marine Sciences, 74: 2223–2238*

7) Has the WG contributed to Advisory needs? If so, please list when, to whom, and what was the essence of the advice.

Not specifically.

8) Please list any specific outreach activities of the WG outside the ICES network (unless listed in question 6). For example, EC projects directly emanating from the WG discussions, representation of the WG in meetings of outside organizations, contributions to other agencies' activities.

Theme session on measuring and assessing biodiversity was held at ICES ASC 2015, Copenhagen. WGIODIV also decided to meet in a non-ICES country (Italy) to attract the expertise of Mediterranean colleagues.

9) Please indicate what difficulties, if any, have been encountered in achieving the workplan.

Completing the final report after a three-year term was challenging, as the participation of some group members was irregular and coordinating the intersessional work has been proven as difficult. For some topics the meetings did not provide enough time to go through all produced text and hence the full report could not be discussed in plenary. As some members were only attending one meeting of the three-year cycle, it was also not easy to decide whether to keep their contribution in the final report. To widen the scope of WGBIODIV, our work could benefit from the expertise of Canadian and US members.

Future plans

- 10) Does the group think that a continuation of the WG beyond its current term is required? (If yes, please list the reasons)

Yes, because the indicator concepts need to be operationalised and applied to different marine regions. Therefore, the work from previous three-year term has built a valuable foundation for newly proposed ToRs (ToR 1&2). Additionally, a new ToR (ToR3) on implication of spatial management measures has been included.

- 11) If you are not requesting an extension, does the group consider that a new WG is required to further develop the science previously addressed by the existing WG.

Not applicable, because WGBIODIV should be extended.

- 12) What additional expertise would improve the ability of the new (or in case of renewal, existing) WG to fulfil its ToR?

The inclusion on wider expertise on different taxa and marine regions (especially the Mediterranean, Baltic Sea and North America) is desirable. There is still a major focus on benthos and fish communities in the North Sea. Also members with expertise in spatial statistics and conservation planning would be highly welcome.

- 13) Which conclusions/or knowledge acquired of the WG do you think should be used in the Advisory process, if not already used? (please be specific)

Consider how WGBIODIV indicator concepts (and eventually indicator assessments), once further validated and applied in case studies, could be included into the ICES ecosystem overviews.