



Ecological risk assessments to guide decision-making: Methodology matters

Gerjan J. Piet^a, Antony M. Knights^{b,1}, Ruud H. Jongbloed^a, Jacqueline E. Tamis^a, Pepijn de Vries^a, Leonie A. Robinson^b

^a Institute for Marine Resources and Ecosystem Studies (IMARES), Haringkade 1, 1976 CP, IJmuiden, The Netherlands

^b School of Environmental Sciences, University of Liverpool, Nicholson Building, Liverpool, L69 3GP, UK

ARTICLE INFO

Article history:

Received 13 May 2016

Received in revised form 18 November 2016

Accepted 19 November 2016

Available online xxx

Keywords:

Ecosystem-based management

Exposure-effect

Sector

Pressure

Ecosystem component

ABSTRACT

Ecological risk assessment is often applied to guide the decision-making process that underpins ecosystem-based management and prioritisation of risk factors for management. Several studies have recently used ecological risk assessment approaches to identify risk factors of greatest concern, but rarely are the underlying methodological decisions discussed in terms of the effect that those decisions have on the outcome of the assessment and ultimately, how that affects prioritisation of risk factors for management.

This study therefore evaluates the effect of methodological decisions involving (1) the choice and definition of risk factors, and (2) the calculation of risk scores, providing, where possible, recommendations on what should be the most appropriate methodologies.

The definition of risk factors is often determined by the policy context and could result in the comparison of one broadly defined risk meta-factor (e.g. Food Production) with corresponding specific risk factors defined more narrowly (i.e. Oil and Gas production or Offshore Wind). Depending on the method to calculate risk this may result in a systematic bias prioritising any risk meta-factor. For the calculation of individual impact chain risk scores we compared weighted scores with ordinal scores, where the former allows more flexibility to represent the qualitative categories that determine risk and provided results better supported by scientific evidence. A consideration of different risk assessment applications in EBM showed there is no one-size-fits-all solution to this as these methodological decisions need to be considered in concert and the preferred methodology may depend on the context in which the risk assessment is applied. The outcome of the risk assessment should always be accompanied by an explicit consideration of these methodological issues and description of the resulting methodological choices.

© 2016 Published by Elsevier Ltd.

1. Introduction

Large areas of marine ecosystems are currently impacted by human activities and many resources are exploited at an unsustainable rate, (e.g. Glover and Smith, 2003; Halpern et al., 2008). As such, sustainable use has become the central paradigm of many recent environmental policies as well as renewed efforts to identify, manage and limit the impact of human activities (e.g. Halpern and Fujita, 2013; Piet et al., 2015; Knights et al., 2013), but the number of threats and constraints on resources can restrict management to a limited number of options and not necessarily those posing the greatest threat to natural systems (Gibbs and Browman, 2015). Decision-support tools are continually being developed (e.g. Jeffrey, 1983, 1992; Piet et al., 2015; Resnik, 1987; Samhuri and Levin, 2012) to support effective decision-making in light of those constraints. Ecological (or environmental) risk assessment (ERA) is an approach that provides a flexible, problem-solving solution capable

of linking the relationship between human activities and the environment, thereby supporting the decision-making needs of environmental managers (Hope, 2006).

Risk assessment *per se* covers a broad array of approaches for a wide set of applications (see reviews by Holdgate, 1979; Evans, 2004; Fryer et al., 2006) here we focus on the approaches most suited to ecological risk assessments (e.g. Astles et al., 2006; Campbell and Gallagher, 2007; Fletcher, 2005). In general terms, ERA describes the likelihood and consequences of an event and can be used to evaluate the degree to which human activities interfere with the achievement of management objectives (Samhuri and Levin, 2012). In this context, risk can be assessed using quantitative (e.g. Francis and Shotton, 1997; Samhuri and Levin, 2012) or qualitative approaches (e.g. Fletcher, 2005; Breen et al., 2012; Fletcher et al., 2010). Traditionally in ERA the likelihood-consequence approach was used for estimating the risk of a rare or unpredictable event (Williams et al., 2011), but when an assessment of the risks associated with on-going (current) pressure is needed, an exposure-effect analysis is more suitable (Smith et al., 2007). Such approaches have been used to consider the potential for ecosystem-based management (EBM) at sub-regional (Samhuri and Levin, 2012) or regional scales (Piet et al., 2015) by allowing decision makers to explore how different management options could reduce threat to their ecosystem policy objectives across a wide range of risk factors. Environmental risk as-

Corresponding Author at: Wageningen IMARES, P.O. Box 68, 1970 AB IJmuiden, The Netherlands.

Email address: gerjan.piet@wur.nl (G.J. Piet)

¹ Present address: Marine Biology and Ecology Research Centre, School of Marine Science and Engineering, Plymouth University, Drake Circus, Plymouth, PL4 8AA, UK.

assessment concepts have also been used to provide a clear structure for cumulative effects assessment for which, according to (Judd et al., 2015), no consistency or standardisation in approaches exists. Risk assessment is therefore playing an increasingly important role in integrating science, policy and management (CENR, 1999).

Any assessment of risk caused by human activities on an ecosystem will be dependent on (1) a correct description of the functioning ecosystem and how this is impacted by those activities, together with (2) an appropriate methodology to translate the impact into risk. Significant progress has been made toward linking human activities to ecosystem impact with the definition and evaluation of the array of sector-pressure-state combinations or “impact chains”, although the resulting network of interactions can be complex (Knights et al., 2013; Tamis et al., 2016; see illustration of impact chains in Fig. 1). Applying a productivity-susceptibility analysis (e.g. Hobday et al., 2011; Samhour and Levin, 2012; Stobutzki et al., 2001) or an exposure-effect evaluation on an interaction network can enable risk to the ecosystem from a single or combination of (anthropogenic) impacts, to be determined (e.g. Bax and Williams, 2001; Halpern et al., 2007; Knights et al., 2015; Milton, 2001; Stobutzki et al., 2001). If risk represents the cumulative effects of different human activities impacting on multiple ecosystem components through different pressures, then individual impact chains need to be combined into an overall measure of risk such that those risk factors, e.g. sector(s), pressure(s) and ecosystem component(s), introducing the greatest level of risk can be identified (Tamis et al., 2016).

As described above, ERA provides a powerful approach for comparison of the effects of different anthropogenic drivers acting on ecosystems (Gibbs and Browman, 2015), but there are many methodological issues to consider in the design of an ERA (Tamis et al., 2016) and we suggest that there has been little discussion in the academic literature of how the decisions made on methodological design affect the outcomes and the advice that is based on this. For example, the method of combining assessment criteria and/or impact chains can vary between studies, in some instances, calculated as the sum of the impact chain scores (e.g. Fock, 2011; Halpern et al., 2008;

HELCOM, 2010; Korpinen et al., 2012; Samhour and Levin, 2012; Stelzenmuller et al., 2010), in others, by the average of them (e.g. Knights et al., 2015; Samhour and Levin, 2012). The values assigned to each assessment category can also vary widely (Tamis et al., 2016). Methodological decisions are clearly made when designing policy-relevant ecosystem assessments, but in the ecological risk assessment approaches outlined to date, there is often a limited description of how these decisions were reached. For example, when choosing the scoring and summation approach, no consideration is given to how this might affect the prioritisation of threats to marine ecosystems and their management. We argue that this exploration of methodological decisions must be openly undertaken and the implications for prioritisation of management explored so that informed decisions can be made about the design of risk assessment to best fit the context in which it is applied.

Here we explore how the methods used to score individual impact chains and to aggregate impact risk over these chains can affect ERA outcomes in terms of the prioritisation of threats. We do this by taking an existing risk assessment approach (Knights et al., 2015), and examine how the outcome of the risk assessment in terms of the rank order of risk factors is altered by (1) changes in the way individual impact chains (within a risk factor) are scored, (2) the method by which multiple chains are then aggregated for an overall risk factor score, and (3) the number of impact chains included, which is often determined by (4) the choice and definition of those risk factors. We also use a case study to discuss the findings in the context of two different cumulative effects assessment applications (following Judd et al., 2015), focusing on how differences in approach methodology can affect: (i) the identification of the most threatening impact chains (see Knights et al., 2015 and analogous to the identification of a ‘hazard’), and (ii) the evaluation of the performance of management measures applied to reduce the risk from specific impact chains (described in full in Piet et al., 2015).

2. Material and methods

The ERA framework evaluated here was based on a sector-pressure-ecosystem component linkage matrix broadly consistent with the interactions possible in European regional seas (based on White et al., 2013). Each of these interactions (herein referred to as impact chains) had earlier been categorised following the methods outlined in Robinson et al. (2013) using five assessment criteria ((criteria: (1) spatial exposure, (2) temporal exposure, (3) impact/severity where exposure occurs, (4) resilience of affected ecosystem components, and (5) persistence of the pressure in the ecosystem). Each impact chain was given a categorical valuation for each of the five assessment criteria (see 2.1 below); the value derived using expert judgement underpinned by a combination of qualitative and quantitative data through a series of expert workshops. In this study, we focus on what we refer to as the impact risk (IR), which is generated using the assessment of exposure and severity criteria converted into a risk score based on numerical values assigned to each categorical descriptor (see Table 1, and more detail in Knights et al., 2015). A number of descriptive statistics can be used to describe IR for individual or aggregations of impact chains. For example, total IR can be calculated by aggregating the scores of individual impact chains associated with a specific sector, pressure or ecosystem component (see Fig. 1). For example, considering Fig. 1, the total IR associated with the sector across its 12 impact chains could be calculated, or the total risk associated with ecosystem component 2 could be calculated (which would just involve aggregating the scores from the 3 relevant impact chains if restricted to this subset). However, the method by

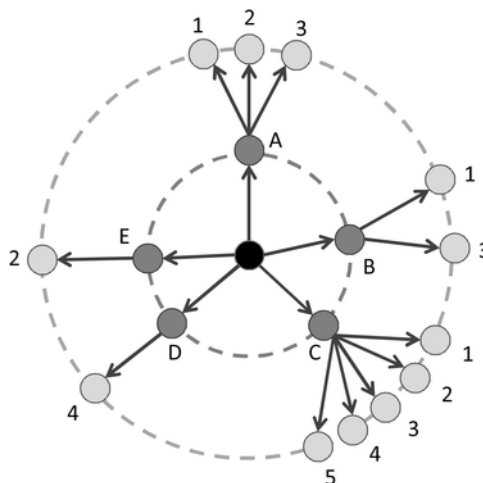


Fig. 1. A schematic representation of the impact chains generated by the activities of a single sector (centre black node), which generates 5 pressures (nodes A–E, inner circle) that impact 5 ecosystem components (nodes 1–5 outer circle). The total number of impact chains generated by the sector is 12. Note that each ecosystem component (EC) can be impacted by several pressures (e.g. EC1 is impacted by pressures A, B and C). Grouping by Sector results in 12 impact chains; Pressure (3, 2, 5, 1 and 1 for respectively pressures A–E); and ecosystem component (3, 3, 3, 2 and 1 for respectively EC1–5). The diagram is modified from the impact chain schematic first shown in Knights et al. (2013).

Table 1

Standardised scores (ordinal and weighted) applied to each impact chain in the risk assessment. Criterion, categories and standardised weighted scores are adapted from Knights et al. (2015) (see justification in Table 1 of Knights et al. (2015), for weighting of scores), whilst the ordinal scores represents an even spacing between categories. Both ordinal and weighted scores are standardised between 0 and 1.

Criterion	Category	Scores	
		Ordinal	Weighted
Spatial Extent (spatial exposure)	Site	0.33	0.03
	Local	0.67	0.37
	Widespread	1	1
Frequency of overlap (temporal exposure)	Rare	0.25	0.08
	Occasional	0.5	0.33
	Common	0.75	0.67
	Persistent	1	1
Degree of Impact (severity of interactions)	Low	0.33	0.01
	Chronic	0.67	0.13
	Acute	1	1

which numerical scores can be aggregated can differ and may affect the ranking and conclusions drawn from the assessment. For example, aggregate values could be calculated using either the SUM or AVERAGE of risk scores.

Below we describe an analysis undertaken to explore how altering the method of score aggregation can affect the outcomes of an ERA. In this instance, we explore this by considering the effect of different aggregation methods on the rank order of impact chains based on their IR scores.

2.1. Scoring of impact chains

For each impact chain, a scoring method was applied to each of the original qualitative assessment categories for the three relevant criteria described above using one of two scales:

- Ordinal – the spacing between categories for each risk criterion is assumed to be equal;
- Weighted – values between categories are weighted to reflect more accurately the content of the categories (e.g. incorporating prior knowledge of the differences between risk criteria categories, see Table 1 in Knights et al., 2015).

Both scoring methods were converted to standardised scales (see Table 1) for the purpose of comparing effects of scoring on the outcomes of the risk assessment overall.

2.2. Method for aggregating risk across impact chains

In the risk assessment, all impact chains are individually scored. Impact chains can then be aggregated depending on interest and context. For example, chains can be grouped by sector (thereby aggregating all pressures introduced by that sector and all ecosystem components those sector-pressure combinations interact with), pressure (aggregating all sectors introducing the pressure type and all ecosystem components those sector-pressure combinations interact with), or ecosystem component (aggregating all sector(s)/pressure(s) combinations affecting a specific ecosystem component) thereby allowing an estimate of total IR to be determined for each combination.

Risk scores from multiple impact chains can be aggregated in several ways. Here, we explore how the use of four different descriptive statistics (i.e. sum (SUM), median (MED), average (AVG) and maximum (MAX)) can affect the rank order of outcomes and thus, our assessment of risk. We used Spearman rank correlation to compare the rank order of the different risk factors, i.e. sectors, pressures and

ecosystem components, depending on the applied statistic. If the rank correlation between two statistics is not significant ($p \geq 0.05$) the application of these statistics will give a different prioritisation of the risk factors considered.

2.3. Numbers of impact chains

Not only does the method of aggregation, but also the number of impact chains, have the potential to affect aggregate risk scores depending on the distribution of risk values. Here, we explore how the number of impact chains can (i) affect the outcome of the risk assessment in terms of total impact risk generated, and (ii) consider how this varies depending on the method of aggregation (see 2.2 above).

2.4. Choice and definition of risk factors

The decision to aggregate impact chains and the focus of the assessment (e.g. sector, pressure or ecosystem components, see Tables 2–4 respectively) determines which and how many impact chains are included in the analysis. Our test involved grouping impact chains firstly by sectors, pressures or ecosystem components, and secondly,

Table 2

The number of impact chains (#) per sector and the ranking of the sectors in terms of their relative importance based on calculated IR which depended on methodological choices: application of weighted (W) instead of ordinal (O) scores, different statistical methods (AVG or SUM), or groupings, i.e. original specific Risk factors (here sectors) versus their corresponding Risk meta-factors (based on EC, 2012). For the Risk meta-factor the ranking is based on categorical scores only distinguishing between a single substitution (S) of the specific Risk factor(s) with their corresponding Risk meta-factor or the simultaneous substitution of all (A) specific sectors with their corresponding Risk meta-factor. NB the Risk meta-factor number of impact chains (#) is not equal to the sum of the # of the corresponding specific Risk factors because in cases where identical impact chains occur in several specific Risk factors (i.e. in case of sector with the same pressure and ecosystem component), then only the impact chain with the highest risk score is retained and all others are removed.

Specific Risk factors: Sector	Risk meta-factor											
	#		AVG		SUM		#		AVG		SUM	
	O	W	O	W	O	W	S	A	S	A	S	A
Nuclear Power	63	15	17	13	15	Energy production	103	8	6	6	6	5
Oil & Gas	86	11	8	4	7	Extraction non-living resources	46	11	7	11	8	
Renewable Energy Aggregates	58	12	14	11	13							
Aquaculture	78	1	5	1	5	Food production	111	1	1	1	1	1
Fishing	76	13	1	7	1	Land-based activities	90	3	3	2	3	
Harvesting/Collecting Agriculture	8	5	12	17	17							
Land-based Industry	56	4	3	5	4	Man-made structures	82	8	5	8	6	
Waste Water Treatment	82	2	6	2	6							
Coastal Infrastructure	38	6	10	12	12	Military Recreation	56	13	8	10	7	
Telecom	71	10	7	6	8							
Military	41	17	16	16	16	Research Transport	57	15	9	14	9	
Tourism/Recreation	56	9	13	9	10							
Research	78	3	4	3	3	Dredging	77	3	2	2	2	
Navigational	57	16	15	15	14							
Shipping	45	8	9	10	9							
	54	14	2	14	2							

Table 3

The number of impact chains (#) per pressure and the ranking of the pressures in terms of their relative importance based on calculated impact risk which depended on methodological choices. For further explanation see Table 2.

Specific Risk factors: Pressure	Risk meta-factor									
	#		AVG		SUM					
	O	W	O	W	O	W				
	S	A	S	A	S	A				
Microbial pathogens	62	8	13	5	8	Biological disturbance	92	1	1	1
Non-indigenous species	30	7	15	11	13					
Selective extraction of species	37	3	1	6	1					
Non-synthetic compounds	114	14	10	2	4	Contamination	135	9	4	2
Radionuclides	35	22	23	16	21					
Synthetic compounds	133	9	9	1	2					
Input of organic matter	30	5	3	9	5	Enrichment	37	3	2	5
N and P enrichment	28	1	4	7	7					
Change in Wave Exposure	6	21	12	22	22	Interference with hydrological processes	57	11	5	10
Emergence regime change	5	23	11	24	23					
pH changes	20	10	8	15	12					
Salinity regime changes	16	11	17	19	19					
Thermal regime changes	18	13	21	18	20					
Water flow rate changes	35	17	20	14	16					
Barrier to species movement	8	2	7	21	18	Other physical disturbance	93	3	3	2
Electromagnetic changes	6	24	24	23	24					
Marine Litter	76	12	5	4	3					
Underwater noise	42	15	22	12	17					
Abrasion	48	16	16	10	11	Physical damage	132	15	7	5
Changes in Siltation	119	18	18	3	6					
Death or injury by collision	15	4	2	17	9					
Selective Extraction (non-living) resources	11	6	6	20	15					
Sealing	56	20	14	8	10	Physical loss	73	14	6	8
Smothering	43	19	19	13	14					

using broader classifications ('meta-groups') based on European Commission groupings such as *Energy Production* or *Transport* (see Table 1; (EC, 2012)).

To test the effect of grouping, we firstly undertook (1) a substitution of specific components for the corresponding risk meta-factor, e.g. two specific sectors *Shipping* and *Navigational dredging* were replaced by *Transport*, or three specific sectors *Nuclear Power*, *Oil & Gas* and *Renewable Energy* were replaced by *Energy production*; followed by (2) the simultaneous and multiple substitution of all the original specific components by their corresponding risk meta-factor. The test applied both AVG and SUM to aggregate across impact chains.

Table 4

The number of impact chains (#) per ecosystem component and the ranking of the ecosystem components in terms of their relative importance based on calculated impact risk which depended on methodological choices. For further explanation see Table 2.

Specific Risk factors: Ecosystem component	Risk meta-factor										
	#		AVG		SUM						
	O	W	O	W	O	W					
	S	A	S	A	S	A					
Deep-sea fish	21	9	1	11	9	Fish	131	2	2	1	2
Demersal fish	130	2	3	1	2						
Pelagic fish	130	4	6	2	3	Marine mammals and reptiles	109	10	4	6	4
Marine mammals and reptiles	109	8	10	5	6	Seabirds	79	11	5	10	5
Seabirds	79	11	11	9	10	Seafloor habitat	138	2	1	1	1
Deep sea bed	28	5	4	10	11						
Littoral rock	92	7	8	7	7						
Littoral sediment	114	6	7	4	4						
Sublittoral rock	90	10	9	8	8						
Sublittoral sediment	126	3	2	3	1						
Water Column	74	1	5	6	5	Water Column	74	5	3	5	3

3. Results

The effect of the methodological decisions on the outcome of this ERA in terms of the main risk factors, i.e. sectors (Table 2), pressures (Table 3), or ecosystem components (Table 4), was based on a comparison of their rank order.

3.1. Scoring of impact chains

Using weighted instead of ordinal scores had a considerable effect on the risk assessment in terms of the relative importance (ranking) of impact chains (see Tables 2–4). For example, *Fishing* and *Shipping* were identified as the greatest risk factors using weighted values, but *Aquaculture* and *Land-based industry* were of greatest risk when ordinal values were used (Table 2). Similarly, there was a change in priority for pressures (Table 3) and ecosystem components (Table 4). For pressures, *Selective extraction* was identified as the greatest risk factor using weighted values, whereas *Nitrogen and Phosphorus enrichment* was the greatest risk factor when using an ordinal scale, and for ecosystem components, greatest risk was to *Deep-sea fish* and the *Water column* respectively.

3.2. Method for aggregating risk across impact chains

The choice of descriptive statistic to describe risk greatly affected the relative importance of each sector, pressure or ecosystem component (Table 5). This effect was most obvious for ecosystem components where the highest ranking component(s) and all interactions were non-significant ($p \geq 0.05$) with a rank correlation varying between -0.07 (MED-MAX) to 0.34 (SUM-MAX) except for the interaction AVG-MED with a rank correlation of 0.79 . To give an example comparing the statistical methods most often applied: SUM would identify *Sublittoral sediment* as the ecosystem component mostly at risk while AVG would identify *Deep sea fish* (see Table 4).

Table 5

Assessment of the effect of the statistical aggregation method on the relative importance in terms of impact risk based of the sector, pressure and ecosystem components using Spearman rank correlation. Indicated is the correlation coefficient, in bold if not significant ($p \geq 0.05$). The correlation coefficient AVG-SUM corresponds to the ranks shown in Tables 2–4 for weighted scores (i.e. Specific risk factors-W column).

		AVG	MED	MAX
Sector	SUM	0.94	0.85	0.84
	AVG	–	0.91	0.86
	MED	–	–	0.72
	MAX	–	–	–
Pressure	SUM	0.61	0.71	0.65
	AVG	–	0.82	0.44
	MED	–	–	0.64
	MAX	–	–	–
Ecosystem component	SUM	0.31	0.22	0.34
	AVG	–	0.79	0.25
	MED	–	–	-0.07
	MAX	–	–	–

3.3. Numbers of impact chains

The number of impact chains was a determinant of ranking for sector, pressure or ecosystem component groupings (Fig. 2, Tables 2–4). In general, the highest risk impact chains are always likely to be included in a risk assessment, whereas lower risk impact chains may not be included (although this decision may not be underpinned by evidence). Increasing the number of impact chains will increase the risk if the SUM method is applied, but decrease the risk if the AVG is applied. This because the risk score for additional impact chains is likely to be low (Fig. 2). There are advantages and disadvantages to both descriptive statistics. An advantage of the SUM method is that risk approaches the asymptote after 25 of the most important impact chains are included and is therefore relatively robust to the addition of new (and relatively less important) impact chains (Fig. 2a). In contrast, risk estimates continue to decline under the AVG method with the addition of new chains (Fig. 2b), implying it is more sensitive to the inclusion of additional impact chains and aggregate assessments with many (relatively unimportant) impact chains may suggest the threats are of low risk.

The effect of aggregation method can easily be demonstrated using examples with many or few impact chains. Take the two ecosystem components with the highest number of impact chains (*Demersal fish* and *Pelagic fish* with 130 chains each, see Table 4); these are ranked #1 and #3 when based on the 25 highest risk chains only. However, including all 130 changes in the analysis decreases their rank (importance) to #3 and #6 respectively. In contrast, for ecosystem components with relatively few impact chains (e.g. *Deep sea fish* with 21 chains and deep seabed with 28 chains), their rank increases from #5 to #1 and #11 to #4 respectively when the risk is calculated using the 25 most important chains or all chains. This effect is less apparent when aggregating risk by sector and pressure as (i) the number of impact chains are generally more evenly distributed among the sectors, and (ii) the pressures with many impact chains have relatively low rankings irrespective of the number of chains included. Thus, when the aggregated risk score is based on AVG, the relative importance of a specific sector, pressure or ecosystem component may heavily depend on the number of impact chains included in the analysis and act in a similar manner to log transformation of biodiversity data to account for rare species.

3.4. Choice and definition of risk factors

The assessment of the relative importance of sectors, pressures and ecosystem components shows that, without exception, the single substitution of the specific risk factor(s) (i.e. sectors, pressures or ecosystem components) for their corresponding risk meta-factor (see Tables 2–4) would always increase the importance of the risk meta-factor if based on SUM. One example based on sectors (Table 2) shows that *Land-based activities* become the 3rd most important sector in terms of its contribution to impact risk (Risk meta-factor SUM-W-A) while the specific sectors *Agriculture*, *Land-based Industry* and *Waste Water Treatment* would only rank 4, 6 and 12 respectively (Specific Risk factors SUM-W-S). The other methods (only AVG shown but this also applies for MED and MAX) would usually result in the ranking of the risk meta-factor within the boundaries of the corresponding specific risk factors, equal (MAX) or close (MED more so than AVG) to that of the worst-case specific risk factor. If the risk assessment is based on risk meta-factors only (see Table 2 column Risk meta-factor-SUM-W-A) the rank, and thus importance, of each risk meta-factor is very similar to the relative importance of the corresponding worst-case specific risk factor (see Table 2 column Specific Risk factors-SUM-W). The definition and choice of risk factors that determines which and how many impact chains are aggregated into the risk factor is therefore mostly relevant when the SUM method is applied. In that case, however, it can systematically bias the outcome of the assessment towards an increased importance of the risk factors with most impact chains (i.e. Risk meta-factors).

4. Discussion

In this study we explored how methodological decisions in the design of a risk assessment may determine the outcome of this risk assessment thereby compromising its ability to guide the decision-making process toward EBM (Fletcher, 2005). We evaluated the effect of methodological decisions involving (1) changes in the way individual impact chains (within a risk factor) are scored, (2) the method by which multiple chains are then aggregated for an overall risk factor score, and (3) the number of impact chains included, which is often determined by (4) the choice and definition of those risk factors. The outcome of this analysis showed that each of the methodological issues considered have the potential to affect the outcome of the risk assessment in terms of a change in the prioritisation of the risk factors (i.e. sectors, pressures or ecosystem components). In order for risk assessments to inform policy and guide managers to address a specific environmental issue, it is necessary to identify appropriate risk factors determined by the policy context (Piet et al., 2015) for which their contribution to perceived risk guides management towards those factors contributing most to that risk (e.g. Fletcher, 2005; Francis, 1992; Hobday et al., 2011; Samhuri and Levin, 2012; Smith et al., 2007). The choice and definition of risk factors (represented by the nodes in Fig. 1) is similar to the issue of what should be considered appropriate “basic elements” (Tamis et al., 2016) but also includes how these are connected in terms of the number of linkages (i.e. impact chains in Fig. 1). Following Piet et al. (2015), our prioritisation was based on the ranked order and therefore the assessed methodologies are only considered to affect the outcome of the risk assessment if the ranked order itself is affected, a much stricter criterion than a change in calculated risk scores. In this study, the risk assessment aggregates risk across impact chains into different risk factors and prioritises these risk factors in terms of their like-

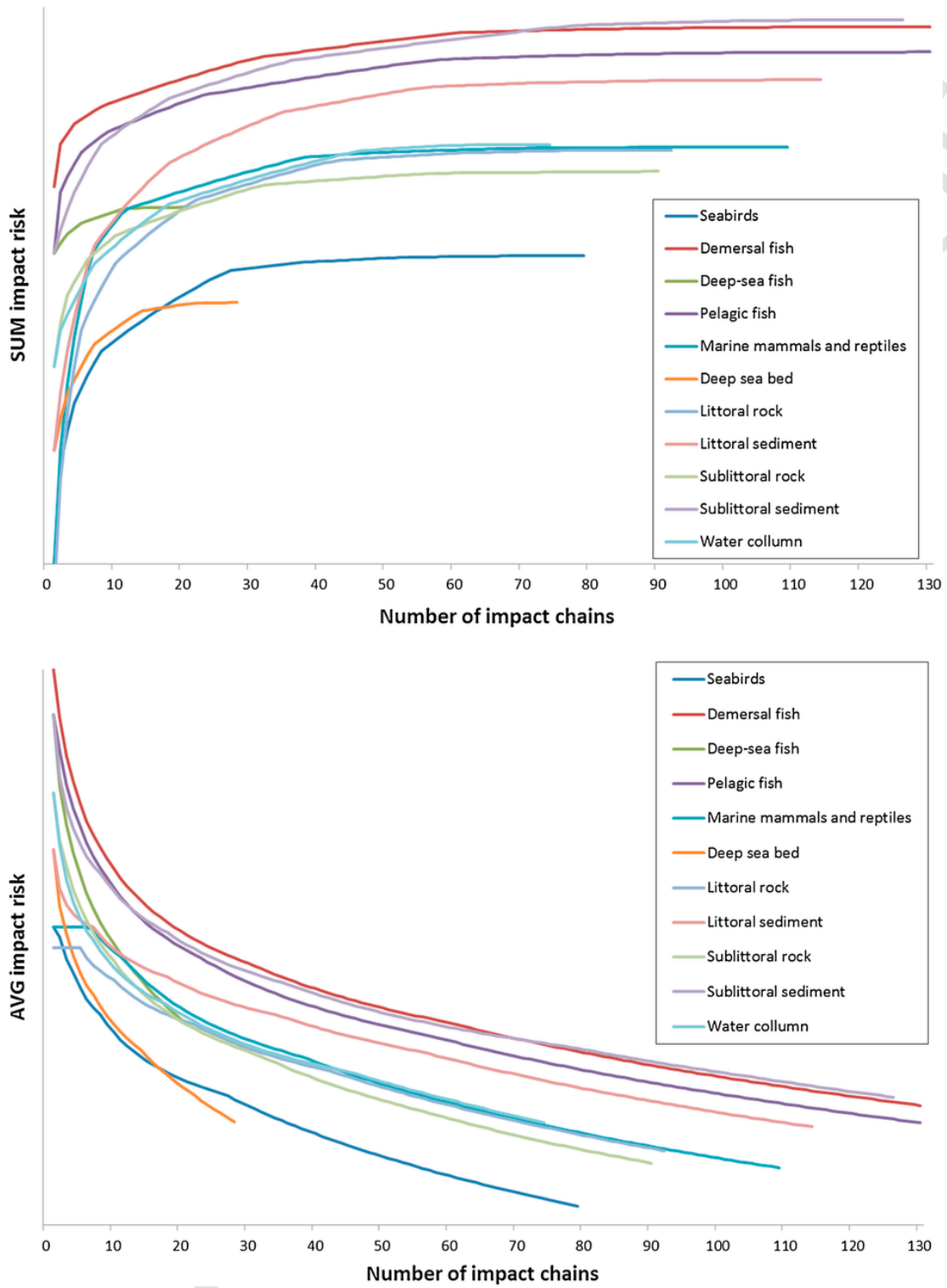


Fig. 2. The impact risk score per ecosystem component based on the summed (SUM) and average (AVG) impact risk depending on the number of impact chains with the importance of the impact chains in terms of their contribution to risk decreasing from left to right. At each specific number of impact chains the rank is determined by the vertical cross section decreasing from top to bottom.

likelihood to cause an adverse impact on the marine ecosystem (Bottrill et al., 2008; Knights et al., 2014).

4.1. Scoring of impact chains

The effect of the scoring of impact chains was tested by comparing the application of weighted scores instead of ordinal scores where the former was believed to better represent the criteria that determined risk (see Table 1). This showed that the choice of risk scores affects the prioritisation of risk factors (i.e. sectors, pressures or ecosystem components) in terms of their contribution to risk. We then evaluated the different outcomes in order to determine which of the two scoring methods performs best in terms of delivering results that are supported by scientific evidence. The application of weighted scores instead of the often applied ordinal scores (e.g. Samhoury and Levin, 2012) revealed *Fisheries* and *Selective extraction of species* as respectively the most important sector and pressure in terms of their contributions to the risk of causing an adverse impact on the ecosystem. This agrees with the statement that fishing is the most widespread human exploitative activity in the marine environment and mainly responsible for the adverse impacts, direct or indirect, on the ecosystem (Jackson et al., 2001; Jennings and Kaiser, 1998). The ordinal scores showed *Agriculture* to be the most important sector while *Nitrogen and Phosphorus enrichment* came out as the main pressure, both results which are not supported by any scientific evidence. These findings suggest that the weighted scores – which are assumed to represent the actual content of the categories better – indeed appear to improve the performance of the risk assessment.

4.2. Method for aggregating risk across impact chains

The method applied for the aggregation of risk scores across impact chains was tested using four different descriptive statistics, namely sum (SUM), median (MED), average (AVG) and maximum (MAX) as aggregation methods (see Table 5). The methodology affected the ranking and thus, our prioritisation of risk factors. Which of these aggregation methods is most appropriate is difficult to assess. When applying weighted scores, both average and sum aggregation methods would prioritise the same sectors and pressures – the outcomes of which can be justified and supported by current scientific evidence. However, the use of averaging (AVG) was found to be more sensitive to the number of impact chains (see 4.3) and the choice and definition of risk factors (see 4.4). Other potential aggregation methods (e.g. median and maximum, see Knights et al., 2015), behave similarly to the AVG method. However, while the use of the median descriptive statistic is expected to be least affected by grouping, this may not necessarily be the case due to the effect of removing identical chains (i.e. with the same pressure and ecosystem component in case sectors are aggregated, see Tables 2–4 for how much chains are removed). Using MAX will always deliver the worst case aggregated risk score and hence a risk assessment that uses MAX is unlikely to be affected by the choice and definition of risk factors, mostly because the method results in many risk factors with identical (i.e. maximum) risk scores. This also implies that MAX performs poorly in prioritising between the risk factors based on these aggregated risk scores and is therefore not a useful aggregation method. Alternatively, if summation of risk scores is the preferred method, risk factors should be identified in relation to the specific issue the risk assessment is expected to address at the initial stage through e.g. a stakeholder consultation process. If a risk assessment is devised for general application, such as to inform decision-makers on issues related to different policy frameworks for which different risk

factors may apply, it is advisable to initially use the most detailed basic elements, i.e. least aggregated risk factors, allowing best available evidence to be clearly linked to specific impact chains and thus, increase transparency of the decision-making process. Best practice would be to start with those risk factors and include only the most relevant impact chains such that the impact chains are more or less evenly distributed among the risk factors.

4.3. Numbers of impact chains

When considering the number of linkages, we found that irrespective of methodological decision, increasing the number of impact chains will create systematic bias by either (i) increasing the risk if the SUM method is applied, or (ii) decreasing the risk if the AVG is applied (Fig. 2). In general the AVG method appears to be more sensitive to the inclusion of additional impact chains and adding more (spurious) impact chains will systematically lower the ranking of a risk factor (e.g. sector) and hence the likelihood it will be targeted by management action.

4.4. Choice and definition of risk factors

The effect of definition of risk factors was tested using two levels of aggregation: the level of what we called specific risk factors, i.e. sectors, pressures or ecosystem components, and the risk meta-factors consisting of a combination of one or more of its corresponding specific risk factors. Both levels were based on their occurrence and definition in actual policy documents (EC, 2012; EC, 2008). Results show that the choice and definition of risk factors can markedly affect the outcome of the risk assessment, certainly if this affects the number of impact chains belonging to a risk factor, but depending on the aggregation method. If aggregated risk was calculated as the sum of the risk across impact chains (e.g. Fock, 2011; Halpern et al., 2008; HELCOM, 2010; Korpinen et al., 2012; Stelzenmuller et al., 2010) the merging of specific risk factors into larger risk meta-factors will always result in a markedly higher rank of the risk meta-factor than any of its corresponding specific risk factors. In case other statistical methods are applied, i.e. average (e.g. Knights et al., 2015; Samhoury and Levin, 2012), median or maximum, the ranking of the risk meta-factor is expected to fall within the boundaries of the corresponding specific risk factors close to that of the worst-case specific risk factor. This sensitivity to the choice and definition of risk factors in case of summation is not a desirable property of any risk assessment and should therefore best be avoided if there is a chance specific risk factors are going to be merged into larger risk factors (i.e. consisting of more impact chains) such as the risk meta-factors applied in this study. To resolve this, the level of detail of these risk factors should be balanced in terms of the number of impact chains per risk factor (i.e. sector-, pressure- or ecosystem component) and this number should be reported together with the aggregated risk. Artificially creating larger risk factors or deconstructing them into smaller risk factors will cause systematic bias respectively increasing or decreasing their relative importance and can therefore be used to manipulate the outcome of the risk assessment.

4.5. A practical application of ERA as part of EBM

The findings above are here considered using real-world examples of how ERA might be used in an EBM context. We consider two applications: (1) an ERA to determine the main anthropogenic pressure on the ecosystem, and (2) integration of the outcomes of this ERA to advise on management options within an EBM framework in an ef-

fort to improve state and meet specific policy conservation objectives.

For the first application, we consider pressures as the starting point as these are often the focus of EBM (e.g. Nitrates Directive). Using the ERA, we identify the three highest risk pressures using the recommended weighted scores rather than ordinal scores (see Section 4.1) and consider the effect of using AVG or SUM to aggregate risk groupings. Irrespective of the aggregation method, *Selective extraction of species* is the highest risk pressure. However, for the identification of the 2nd and 3rd highest risk pressures, the choice of aggregation method does matter. Using AVG highlights *Death or injury by collision* and *Input of organic matter*, whereas SUM indicates *Introduction of Synthetic compounds* and *Marine Litter* as the 2nd and 3rd highest risk pressures (see Table 3).

Integrating this ERA outcome within an EBM framework, which according to (Judd et al., 2015) should be a key focus of future research), we distinguish between selection and evaluation of EBM options. For the initial selection, both risk aggregation methods (AVG or SUM) point toward a requirement for (additional) fisheries management as the pressure *Selective extraction of species* is almost entirely caused by the *Fishing* sector. However, other management measure options are affected by the methodological choice. Using AVG as applied by Knights et al. (2015) and Halpern et al. (2007), would point toward mitigation of the effects of (i) *Shipping on Marine mammals*, (ii) *Renewable energy* (i.e. windfarms) on *Seabirds*, and (iii) the output of *Organic matter from Agriculture*. In contrast, the SUM method would recommend mitigation of (i) the *Introduction of synthetic compounds* by sectors such as *Oil & Gas* or *Land-based industry*, and (ii) dumping of *Marine litter* by maritime activities. Hence, depending on the methodology, an entirely different suite of management measures would be recommended.

In the next step where the ERA is applied to evaluate the management option(s), we focus in this example only on the highest risk pressure: *Selective extraction of species by Fishing*. We propose specific management measures for the conservation of the main ecosystem components at risk. In this instance, the aggregation method is not problematic (see Section 4.2) and following the recommendation of Piet et al. (2015), we apply the SUM method to find three ecosystem components most at risk namely: *Demersal fish*, *Pelagic fish* and *Sublittoral habitat* (see Table 4, SUM-W column). The evaluation of the management options would then be based on their performance towards achieving the conservation objectives, which in practice, is measured as a reduction in risk to these three ecosystem components. For this example, we selected from a broader list of possible management options (see Piet et al., 2015) three typical fisheries management measures consisting of the use of:

- Total Allowable Catch (TAC) or Quotas that target all fish, i.e. Demersal and Pelagic fish;
- Spatio-temporal closure of the pelagic fishery to target only the pelagic fish; and
- Spatio-temporal closure of the demersal fishery to target demersal fish and seafloor habitat.

How methodological choice alters the outcome of this evaluation process is illustrated by considering two methodological issues, namely: (1) how substitution of some specific risk factors (i.e. the ecosystem components *Demersal fish* and/or *Pelagic fish*) for their meta-factor (*Fish*) changes the rank order of the most at risk ecosystem components, and (2) the effect of including the top 25 impact chains by risk as opposed to including all impact chains. In the first instance, *Sublittoral sediment* is the most at-risk ecosystem component when we consider ecosystem components separately, but

using meta-groupings indicates *Fish* rather than *Habitats* are most at risk (see Table 4). Changing the number of impact chains included changes the focus from *Demersal fish* when using the top 25 impact chains toward *Sublittoral sediments* when all impact chains are included (see Fig. 2a)

Because the management options were shown to differ in terms of the ecosystem components they target, it is very likely that these methodological choices that affect the rank order of these ecosystem components will result in different outcomes of this evaluation and hence different perceptions as to what should be the preferred management option(s) based on performance.

4.6. Conclusion

The implications of this methodological study highlight important considerations relevant to the initial design stage of a new risk assessment, such as one intended to address a specific policy issue or inform EBM, as well as to later stages where an existing risk assessment is being used or adapted to inform on new policy issues or management measures where it was not necessarily designed for. At the initial design stage, the choice and definition of risk factors (e.g. based on DPSIR categories) that comprehensively characterize all the pathways through which human activities impact the ecosystem need to be considered together with appropriate scores for the exposure-effect categories that determine risk for each single pathway. This selection of the risk factors to include in the ERA is directly related to the recommendation by Judd et al. (2015) that any application of an ERA should be accompanied by a detailed description of the problem formulation.

The choice of an appropriate aggregation method then depends on which phase of EBM the risk assessment is used for distinguishing between the *a priori* selection of appropriate management measures and the *a posteriori* evaluation of their performance. The former is illustrated by Knights et al. (2015), who provide box plots showing minimum, median and maximum risk per sector, pressure and ecosystem component, thereby capturing all the relevant aspects of risk. The latter is illustrated by Piet et al. (2015) who choose summed risk for an evaluation of the performance of management strategies to reduce risk, which also appears appropriate as generic measures (e.g. those involving several sectors) are likely to reduce more risk than measures targeting a single sector (or pressure, or ecosystem component).

Thus while this paper has established how the methodology to calculate risk affects the outcome of the ERA, it is clear that there is no one-size-fits-all solution as the choice of what can be considered the best methodology often depends on the context in which the ERA is applied. For transparency in the advisory process any application of an ERA to inform decision-making a minimum requirement is therefore to explicitly consider these methodological issues and report on the resulting methodological choices.

Acknowledgements

This study was funded by the EU FP7 programme 'Options for Delivering Ecosystem-based Marine Management' (ODEMM; grant number 244273; www.liv.ac.uk/odemmm).

References

- Astles, K.L., Holloway, M.G., Steffe, A., Green, M., Ganassin, C., Gibbs, P.J., 2006. An ecological method for qualitative risk assessment and its use in the management of fisheries in New South Wales, Australia. *Fish. Res.* 82, 290–303.

- Bottrill, M.C., Joseph, L.N., Carwardine, J., Bode, M., Cook, C.N., Game, E.T., Grantham, H., et al., 2008. Is conservation triage just smart decision making? *Trends Ecol. Evol.* 23, 649–654.
- Breen, P., Robinson, L.A., Rogers, S.I., Knights, A.M., Piet, G., Churilova, T., Margonski, P., et al., 2012. An environmental assessment of risk in achieving good environmental status to support regional prioritisation of management in Europe. *Mar. Policy* 36, 1033–1043.
- CENR, 1999. Ecological risk assessment in the federal government.
- Campbell, M.L., Gallagher, C., 2007. Assessing the relative effects of fishing on the New Zealand marine environment through risk analysis. *ICES J. Mar. Sci.* 64, 256–270.
- EC, 2008. Establishing a framework for community action in the field of marine environmental policy (Marine Strategy Framework Directive). 2008/56/EC. 40 pp.
- EC, 2012. Guidance for 2012 reporting under the Marine Strategy Framework Directive, using the MSFD database tool. Version 1.0. 164 pp.
- Evans, E., 2004. Foresight: future flooding: scientific summary. *Future Risks and Their Drivers*. vol. I. Office of Science and Technology, London, 366.
- Fletcher, W.J., Shaw, J., Metcalf, S.J., Gaughan, D.J., 2010. An Ecosystem Based Fisheries Management framework: the efficient, regional-level planning tool for management agencies. *Mar. Policy* 34, 1226–1238.
- Fletcher, W.J., 2005. The application of qualitative risk assessment methodology to prioritize issues for fisheries management. *ICES J. Mar. Sci.* 62, 1576–1587.
- Fock, H., 2011. Integrating multiple pressures at different spatial and temporal scales: a concept for relative ecological risk assessment in the European marine environment. *Hum. Ecol. Risk Assess.* 17, 187–211.
- Francis, R. I. C. 1992. Use of risk analysis to assess fishery management strategies - a case-study using orange roughy (*Hoplostethus atlanticus*) on the Chatham Rise, New Zealand. *Canadian Journal of Fisheries and Aquatic Sciences*, 49: 922–930. Francis, R.I.C.C., Shotton, R., 1997. 'Risk' in fisheries management: a review. *Can. J. Fish. Aquat. Sci.* 54, 1699–1715.
- Fryer, M., Collins, C.D., Ferrier, H., Colvile, R.N., Nieuwenhuijsen, M.J., 2006. Human exposure modelling for chemical risk assessment: a review of current approaches and research and policy implications. *Environ. Sci. Policy* 9 (3), 261–274.
- Gibbs, M.T., Browman, H.I., 2015. Risk assessment and risk management: a primer for marine scientists. *ICES J. Mar. Sci.* 72, 992–996.
- Glover, A.G., Smith, C.R., 2003. The deep-sea floor ecosystem: current status and prospects of anthropogenic change by the year 2025. *Environ. Conserv.* 30, 219–241.
- HELCOM, 2010. Ecosystem health of the Baltic Sea 2003-2007: HELCOM initial holistic assessment. 122. 68 pp.
- Halpern, B.S., Fujita, R., 2013. Assumptions, challenges, and future directions in cumulative impact analysis. *Ecosphere* 4.
- Halpern, B.S., Selkoe, K.A., Micheli, F., Kappel, C.V., 2007. Evaluating and ranking the vulnerability of global marine ecosystems to anthropogenic threats. *Conserv. Biol.* 21, 1301–1315.
- Halpern, B.S., Walbridge, S., Selkoe, K.A., Kappel, C.V., Micheli, F., D'Agrosa, C., Bruno, J.F., et al., 2008. A global map of human impact on marine ecosystems. *Science* 319, 948–952.
- Hobday, A.J., Smith, A.D.M., Stobutzki, I.C., Bulman, C., Daley, R., Dambacher, J.M., Deng, R.A., et al., 2011. Ecological risk assessment for the effects of fishing. *Fish. Res.* 108, 372–384.
- Holdgate, M.W., 1979. *A Perspective of Environmental Pollution*. Cambridge University Press.
- Hope, B.K., 2006. An examination of ecological risk assessment and management practices. *Environ. Int.* 32, 983–995.
- Jackson, J.B.C., Kirby, M.X., Berger, W.H., Bjorndal, K.A., Botsford, L.W., Bourque, B.J., Bradbury, R.H., et al., 2001. Historical overfishing and the recent collapse of coastal ecosystems. *Science* 293, 629–638.
- Jeffrey, R.C., 1992. *Probability and the Art of Judgement*. University of Chicago Press, Chicago, Illinois USA.
- Jeffrey, R.C., 1983. *The Logic of Decision*. University of Chicago Press, Chicago, Illinois USA.
- Jennings, S., Kaiser, M.J., 1998. The effects of fishing on marine ecosystems. *Adv. Mar. Biol.* 34, 201–352.
- Judd, A.D., Backhaus, T., Goodsir, F., 2015. An effective set of principles for practical implementation of marine cumulative effects assessment. *Environ. Sci. Policy* 54, 254–262.
- Knights, A.M., Koss, R.S., Robinson, L.A., 2013. Identifying common pressure pathways from a complex network of human activities to support ecosystem-based management. *Ecol. Appl.* 23, 755–765.
- Knights, A.M., Culhane, F., Hussain, S.S., Papadopoulou, K.N., Piet, G.J., Raakaer, J., Rogers, S.I., et al., 2014. A step-wise process of decision-making under uncertainty when implementing environmental policy. *Environ. Sci. Policy* 39, 56–64.
- Knights, A.M., Piet, G.J., Jongbloed, R., Tamis, J.E., Robinson, L.A., et al., 2015. An exposure-effect approach for evaluating ecosystem-wide risks from human activities. *ICES J. Mar. Sci.* 72, 1105–1115. <http://dx.doi.org/10.1093/icesjms/fsu245>.
- Korpinen, S., Meski, L., Andersen, J.H., Laamanen, M., 2012. Human pressures and their potential impact on the Baltic Sea ecosystem. *Ecol. Indic.* 15, 105–114.
- Milton, D.A., 2001. Assessing the susceptibility to fishing of populations of rare trawl bycatch: sea snakes caught by Australia's Northern Prawn Fishery. *Biol. Conserv.* 101, 281–290.
- Piet, G.J., Jongbloed, R.H., Knights, A.M., Tamis, J.E., Pajmans, A.J., van der Sluis, M.T., de Vries, P., Robinson, L.A., 2015. Evaluation of ecosystem-based marine management strategies based on risk assessment. *Biol. Conserv.* <http://dx.doi.org/10.1016/j.biocon.2015.03.011>.
- Resnik, M.D., 1987. *Choices: An Introduction to Decision Theory*. University of Minnesota Press, Minneapolis, Minnesota USA.
- Robinson, L. A., White, L., Culhane, F., and Knights, A. M. 2013. ODEMM Pressure Assessment Userguide V.2. ODEMM Guidance Document Series No. 4. University of Liverpool, Liverpool. 15 pp. Samhuri, J., Levin, P.S., 2012. Linking land and sea-based activities to risk in coastal ecosystems. *Biol. Conserv.* 145, 118–129.
- Smith, A.D.M., Fulton, E.J., Hobday, A.J., Smith, D.C., Shoulder, P., 2007. Scientific tools to support the practical implementation of ecosystem-based fisheries management. *ICES J. Mar. Sci.* 64, 633–639.
- Stelzenmuller, V., Lee, J., South, A., Rogers, S.I., 2010. Quantifying cumulative impacts of human pressures on the marine environment: a geospatial modelling framework. *Mar. Ecol. Progress Ser.* 398, 19–32.
- Stobutzki, I.C., Miller, M.J., Jones, P., Salini, J.P., 2001. Bycatch diversity and variation in a tropical Australian penaeid fishery; the implications for monitoring. *Fish. Res.* 53, 283–301.
- Tamis, J.E., de Vries, P., Jongbloed, R.H., Lagerveld, S., Jak, R.G., Karman, C.C., van der Wal, J.T., Slijkerman, D.M.E., Klok, C., 2016. Toward a harmonized approach for environmental assessment of human activities in the marine environment. *Integr. Environ. Assess. Manag.* 9999, 1–11.
- White, L. J., Koss, R. S., Knights, A. M., Eriksson, A., and Robinson, L. A. 2013. ODEMM Linkage Framework Userguide (Version 2). University of Liverpool, Liverpool. Williams, A., Dowdney, J., Smith, A.D.M., Hobday, A.J., Fuller, M., 2011. Evaluating impacts of fishing on benthic habitats: a risk assessment framework applied to Australian fisheries. *Fish. Res.* 112, 154–167.