Ecological Indicators 84 (2018) 619–631

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Original articles

A quantitative framework for selecting and validating food web indicators

Saskia A. Otto[a,b,*], Martina Kadin[b,c], Michele Casini[d], Maria A. Torres[e,f], Thorsten Blenckner[b]

[a] Institute for Hydrobiology and Fisheries Science, Center for Earth System Research and Sustainability, University of Hamburg, Grosse Elbstrasse 133, D-22767 Hamburg, Germany
[b] Stockholm Resilience Centre, Stockholm University, Kräftriket 2B, SE-106 91 Stockholm, Sweden
[c] Current address: School of Aquatic and Fishery Sciences, University of Washington, 1122 NE Boat St, Box 355020 Seattle, WA 98195, USA
[d] Swedish University of Agricultural Sciences, Department of Aquatic Resources, Institute of Marine Research, Turistgatan 5, SE-453 30 Lysekil, Sweden
[e] Swedish University of Agricultural Sciences, Department of Aquatic Resources, Institute of Coastal Research, Skolgatan 6, SE-742 42 Öregrund, Sweden
[f] Current address: Centre of Marine Sciences, University of the Algarve, Campus de Gambelas, 8005-139 Faro, Portugal

ARTICLE INFO

Keywords:
Indicator-testing framework
Performance criteria
Marine strategy framework directive (MSFD)
Generalized additive modelling (GAM)
State space approach
Marine food web indicators
Baltic sea
Zooplankton
Pelagic fish

ABSTRACT

Finding suitable state indicators is challenging and cumbersome in stochastic and complex ecological systems. Typically, a great focus is given to criteria such as data availability, scientific basis, or measurability. Features associated with the indicator's performance such as sensitivity or robustness are often neglected due to the lack of quantitative validation tools. In this paper, we present a simple but flexible framework for selecting and validating the performance of food web indicators. In specific, we suggest a 7-step process in which indicator performances at a regional scale are quantified and visualized allowing for the selection of complementary indicator suites. We demonstrate its application by comparing the performance of pelagic food web indicators for three basins of the Baltic Sea and by assessing the food web status based on selected indicator suites. Our analysis sheds light on spatial differences in indicator performances with respect to direct and indirect pressures, the role of non-linearity and non-additivity in pressure responses, as well as relationships between indicators caused by species interactions. Moreover, our results suggest that the present food web states in the Bornholm and Gotland basins of the Baltic Sea deviate distinctly from an earlier reference period. We advocate the use of our quantitative framework as decision-support tool for selecting suites of complementary indicators under given management schemes such as the EU Marine Strategy Framework Directive.

1. Introduction

The call for ecosystem-based management approaches has widely instigated the use of ecological state indicators. Sets of indicators that represent key ecosystem characteristics and are tightly linked to management objectives are largely needed. As an effective tool these sets of indicators will be increasingly important for ecosystem monitoring and progress assessment in reaching management targets (Levin et al., 2009). Hence, many ecological state indicators have been suggested under national and international policy frameworks such as the United Nations action plan Agenda 21 (UNDESA, 2007). In European Union (EU) marine policy, indicator development has recently progressed as part of the implementation of the Marine Strategy Framework Directive (MSFD) to aid the achievement of Good Environmental Status (GES) of the EU's marine waters by 2020. GES is defined with respect to 11 qualitative descriptors, which describe the environment when GES has been achieved (EU Directive 2008/56/EC). The MSFD descriptor 4 (D4), which focuses on marine food webs, is perhaps the most challenging one since the identification of simple indicators able to assess the status of system with dynamic species interactions and the identification of underlying responses to pressures is difficult (Shephard et al., 2015). Suggested indicators range from single species and species groups to aggregated metrics such as diversity indices, mean community size and trophic level or network indices (Teixeira et al., 2014). While some indicators can be considered as operational, many are still rather conceptual and lack specific assessment benchmarks (Probst and Stelzenmüller, 2015).

Finding suitable food web indicators is generally cumbersome in marine ecosystems as these systems are highly stochastic (Bjørnstad and Grenfell, 2001) and influenced by a multitude of environmental and anthropogenic forcing (Halpern et al., 2008). Furthermore, non-linear and abrupt responses to pressures changes are ubiquitous (Conversi et al., 2015; Möllmann et al., 2015), similarly to other complex systems such as global climate or world economy (Strange, 2007). Non-

* Corresponding author at: Institute for Hydrobiology and Fisheries Science, Center for Earth System Research and Sustainability, University of Hamburg, Grosse Elbstrasse 133, D-22767 Hamburg, Germany.
E-mail address: saskia.otto@uni-hamburg.de (S.A. Otto).

http://dx.doi.org/10.1016/j.ecolind.2017.05.045
Received 7 March 2017; Received in revised form 16 May 2017; Accepted 16 May 2017
1470-160X/

Table 1
Common criteria for selecting environmental state indicators sorted by their type. The five performance criteria relevant for our proposed framework are highlighted in bold.

No	Criterion	References	Type of criteria
1	Type of indicator (state or pressure)	2,3,7,9	Criteria for identifying candidates
2	Meaningful and grounded in research	1,2,4,5,6,7,8	
3	Intelligible and easily interpreted	3,5,6,7	
4	Simple to measure	1,2,6,7,8	
5	Cost-effectiveness	1,4,5,6,7,8,9	
6	Availability of existing (historical) and ongoing data (for reference points, trend and sensitivity analysis)	1,2,3,4,5,6,7,8	
7	Relevant spatial coverage	3,6,7	
8	**Development reflects ecosystem change caused by variation in manageable pressure(s)**	1,2,6,7	Performance criteria
9	**Sensitive or responsive to pressures**	1,2,4,6,8,9	
10	**Robust, i.e. responses in a predictive fashion, and statistically sound**	1,2,6	
11	**Links to management measures (responsiveness and specificity)**	1,2,4,6,7	
12	**Relates where appropriate to other indicators but is not redundant**	2,3,6,7,9	
13	Threshold, reference, or target values established	1,2,3,6,8,9	Beneficial but not a requirement
14	Early warning potential	6,7,8	
15	Applicable in heterogeneous systems, i.e. across a wide set of subsystems	1,6	
16	Application across different management schemes, e.g. different MSFD descriptors	5,7	

1: OECD, 1993; 2: FAO 1997; 3: EEA, 2005; 4: Rice and Rochet, 2005; 5: UNDESA, 2007; 6: Kershner et al., 2011; 7: Shephard et al., 2015; 8: Queirós et al., 2016; 9: Rossberg et al., 2017

stationarity, i.e. spatio-temporal change in the state-pressure relationship (Hunsicker et al., 2016; Stenseth et al., 2004), additionally impedes the development of robust indicators that behave desirably in a consistent and predictable way. Developing a set of indicators hence requires a thorough performance validation, particularly when using empirical data as basis.

Typically, a process for developing indicator sets comprises a specification of management objectives and a first selection of potential indicators based on previously cited research (Siddig et al., 2016) and followed by an evaluation against specific selection criteria (e.g. EEA, 2005; FAO, 1997; Kershner et al., 2011; OECD, 1993; Queirós et al., 2016; Rice and Rochet, 2005). Table 1 lists 16 selection criteria combining the frameworks of earlier studies. Many of these criteria are mainly applied for identifying candidate indicators without requiring data for an evaluation yet (Table 1–Criteria 1–7). All other criteria require time series for their evaluation and either relate to the indicator's performance (Table 1–Criteria 8–12) or are considered as adding information but not necessarily a prerequisite (Table 1–Criteria 13–16). Indicators are usually evaluated either qualitatively or semi-quantitatively based on expert knowledge (e.g.Kershner et al., 2011; Link, 2005; Shephard et al., 2014; Shin et al., 2010a; Tam et al., 2017). Such expert scoring is, however, prone to subjectivity and high variability (Rochet and Rice, 2005), which calls for more quantitative approaches when rating particularly performance criteria (Queirós et al., 2016). The indicator's performance relates ultimately to the quality of the time series data and, thus, a thorough determination of whether the indicator as implemented meets the expected requirements, i.e. a validation, is needed. Unfortunately, frameworks that explicitly assess performance criteria in a quantitative and transparent way do not yet exist (Rossberg et al., 2017).

In the absence of holistic quantitative assessment frameworks, most empirical studies either assess indicator trends and their suitability in reflecting ecosystem changes (e.g. Blanchard et al., 2010; Methratta and Link, 2006; Probst and Stelzenmüller, 2015; Shannon et al., 2010) or the responsiveness of the indicators to pressures and how that links to management (e.g. Fu et al., 2015; Greenstreet and Rogers, 2006; Large et al., 2013; Torres et al., 2017). One important component of the performance criteria is the robustness of the selected indicator. The robustness is evaluated by the predictability of the indicator responses to environmental and anthropogenic pressures (Kershner et al., 2011) and this is usually neglected. Considering the indicator's robustness is however crucial when applying statistical tools to monitoring data with potentially large measurement errors. Statistical bias, i.e. error introduced by applying simple models to complex dynamics, may additionally exist (James et al., 2013). Model-based indicator approaches, in contrast, include more often robustness test by applying e.g. sensitivity analysis (Bourdaud et al., 2016) or test signal approach (Houle et al., 2012) but do not necessarily validate all other performance criteria at the same time. To address these shortcomings, we present a comprehensive, quantitative and transparent framework for validating the performance of indicators and for selecting a robust set of indicators tailored to meet regional conditions and specific management needs.

We demonstrate the applicability of our framework using a set of suggested indicators addressing the MSFD D4 for the Baltic Sea pelagic food web (Fig. 1). The Baltic Sea is a well-studied system that is characterized by strong hydrographic gradients providing different habitats for species, which allows cross-regional comparisons. Major community changes in recent decades have been described (Diekmann and Möllmann, 2010; Möllmann et al., 2009) allowing further for tests of the indicatorś ability to reflect historical development.

In this paper, we first introduce, in Section 2, a refinement of criteria for the performance validation of food web indicators and a scoring scheme. Section 2 further describes statistical tools to assess and

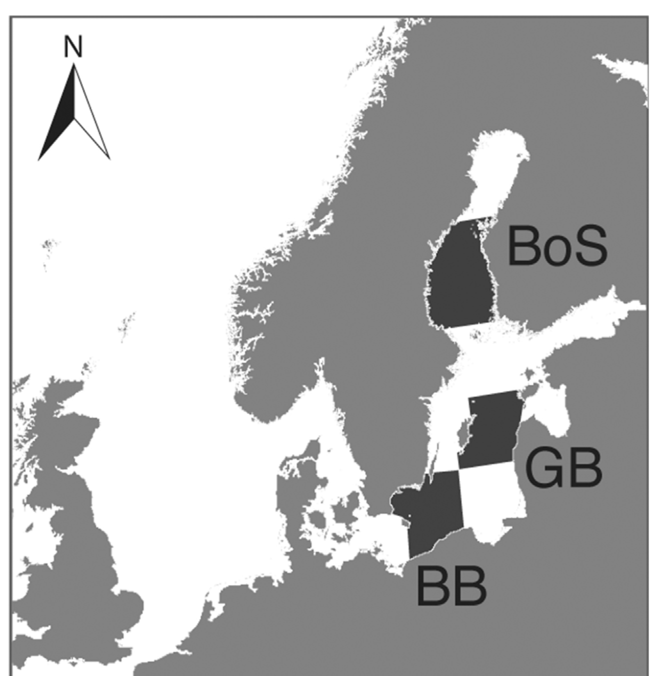

Fig. 1. Map of the Baltic Sea and the location of the 3 study basins (BB = Bornholm Basin, GB = Gotland Basin, BoS = Bothnian Sea).