Contents lists available at ScienceDirect

Ecological Indicators

A simulation-based approach to assess the stability of marine food-webs and inform Good Environmental Status

Joana Brito [a, *], Simone Libralato [b], Telmo Morato [a, c]

[a] *Institute of Marine Sciences - OKEANOS, University of the Azores, 9901-862, Horta, Portugal*
[b] *National Institute of Oceanography and Applied Geophysics-OGS, Section of Oceanography, Trieste, Italy*
[c] *IMAR Instituto do Mar, University of the Azores, 9901-862, Horta, Portugal*

ARTICLE INFO

ABSTRACT

The European Union Marine Strategy Framework Directive (MSFD) recognises that maintaining marine food-webs in Good Environmental Status (GES) is fundamental to ensure the long-term provision of essential ecosystem goods and services. However, operationalising food-web assessments is challenging due to difficulties in i) implementing simple but complete monitoring programmes, ii) identifying thresholds in monitoring indicators that inform when perturbations are diverting food-web state from GES and iii) in providing an integrative and complete picture of the (health) status of food-webs. In this context, stability assessments of marine food-webs could be useful to identifying the indicators that best track perturbation-induced changes in food-web state and the threshold boundaries that should not be exceeded to minimise the likelihood of losing stability. Yet, there is still a lack of systematic methods to perform such assessments. Here, we evaluate the potential of a simulation-based protocol to be used as a methodological standard for assessing the stability of marine food-webs. The protocol draws on the principles of ecological stability theory and provides a framework for assessing the trajectories of individual indicators during perturbation regimes and their robustness in detecting stability thresholds for marine food-webs. We tested the protocol on an open-ocean and deep-sea food-web modelled with the Ecopath with Ecosim suite. We concluded that indicators that quantify transfer efficiency through the food-web and measure the average trophic level of the community are optimal proxies for trophic functioning and structure to assess the stability of the system. Furthermore, we show how the approach can be applied to i) determine the impact of a loss of stability on the balance between trophic levels and ii) identify the biological components of the food-web that are most affected in scenarios of stability loss. Our findings could be useful for the ongoing debate on how trophic models and derived indicators can play a concrete and practical role in the food-web assessments in European seas.

1. Introduction

Several international political agreements call for the implementation of marine ecosystem-based approaches to natural resource management and conservation. In Europe, this effort is guided by the European Union (EU) Marine Strategy Framework Directive (MFSD, Directive 2008/56/EC). The Directive requires EU Member States to implement concrete measures to achieve and maintain Good Environmental Status (GES). Thus, it provides a framework for assessing the integrity, structure and functioning of marine ecosystems to ensure that they remain healthy, productive and capable of providing essential goods and services. GES is to be determined for each of the eleven MSFD descriptors, which qualitatively describe what ecosystems should look like when GES is reached. Among the eleven descriptors, one aims to assess the state of marine food-webs (Descriptor 4, D4), i.e. the networks formed by trophic interactions between species and functional groups in ecological communities (McCann, 2011).

The determination of GES in relation to D4 should be based on specific criteria defined by the MSFD. The two primary criteria assert that marine food-webs are at GES when the effects of (human-induced) perturbations do not adversely affect the diversity (species composition and relative abundance) within and the abundance between trophic guilds (criteria D4C1 and D4C2, respectively) (EU, 2017; European Commission No. 19, 2022). To chart progress towards the achievement

---

* Corresponding author.
E-mail address: joana.cr.brito@uac.pt (J. Brito).

of GES, Member States are expected to implement an integrative indicator-based approach suitable to operationalize these criteria and inform when food-webs reach or deviate from GES using quantitative threshold limits (EU, 2017; European Commission No. 19, 2022). The scientific community has made great efforts to meet the requirements of the MSFD and, therefore, expand current knowledge on food-web dynamics and explore the most appropriate methodological standards for food-web assessments (e.g., Arroyo et al., 2019; Bentley et al., 2019; Kauhala et al., 2019, Korpinen et al., 2022; Otto et al., 2018; Piroddi et al., 2015; Rogers et al., 2010; Rossberg et al., 2017; Rombouts et al., 2013a; Rombouts et al., 2013b; Shephard et al., 2015; Tam et al., 2017; Torres et al., 2017; Tverin et al., 2019). However, the selection of operational indicators and associated thresholds to support the assessment of food-webs needs to be further harmonised and calibrated to overcome key challenges associated to the monitoring of D4 (e.g., Tam et al., 2017; Machado et al., 2021; Shephard et al., 2015). Such challenges include i) difficulties in implementing simple but complete monitoring programmes, ii) difficulties in identifying thresholds in monitoring indicators that inform when perturbations are diverting food-web state from GES, in the absence of reference points describing unimpaired systems (Pauly 1995; Novaglio et al., 2020) and iii) difficulties in providing an integrative and complete picture of the (health) status of food-webs (e.g., Rombouts et al., 2013a; 2013b). In addition, the MSFD is implemented in a six-year cycle that includes reporting, monitoring and review, allowing for improvements, adjustments and continuous adaptation to scientific advances (e.g., Verling et al., 2023).

One possible way to advance the monitoring of D4 is to assess the stability of marine food-webs. Stability captures different facets of food-web dynamics and their responses to perturbations, reflecting the ability of the system to recover from impacts on its state (Donohue et al., 2013, 2016; Grimm & Wissel, 1997; Holling, 1973; Ives & Carpenter, 2007; McCann, 2011; Pimm, 1984; Van Meerbeek et al., 2021). Responses of marine food-webs to perturbation vary according to the specific characteristics of the system and the nature of the perturbation itself (McCann, 2000,2011, Donohue et al., 2013). Perturbations affect food-webs by changing, for example, the number of trophic groups and the connections between the components of the system (e.g. top predators and corresponding prey) (e.g., Bartley et al. (2019), Durante et al. (2022), Kortsch et al. (2015), Ullah et al. (2018)), which can trigger cascading effects and alter trophic structure and functioning (Casini et al., 2009; Link et al., 2015; Möllmann et al., 2015; Nyström et al., 2012). When changes in trophic structure are accompanied by functional changes, recovery to the reference state may take long periods of time (Dubois et al., 2019; Costanza & Mageau, 1999, Moreno-Mateos et al., 2017) or be unlikely because internal feedback mechanisms might stabilize the system in an alternative state (Nyström et al., 2012; Scheffer et al., 2001; Scheffer and Carpenter, 2003; Van Nes & Scheffer, 2007). Therefore, maintaining food-webs within the boundaries of stable states, where they can recover from perturbations and maintain their structure and functioning is crucial for achieving GES, namely criterion D4C2.

A main challenge, however, lies in developing and implementing suitable methods to assess the stability of marine food-webs for monitoring purposes. One promising approach is the use of trophic modelling tools, which quantitatively describe entire food-webs (Christensen and Walters, 2004; Geary et al., 2020). These tools enable the estimation of various indicators that capture the fundamental processes underlying trophodynamic and measure the impact of perturbations on trophic structure and functioning over time (Mageau et al., 1998; Odum, 1985; Ulanowicz, 2004). Tracking the trajectories of these indicators would possibly allow to determine the thresholds values from which the food-web cannot recover from the perturbation's effects (Costanza & Mageau, 1999). However, despite their potential operational value, there is still a lack of systematic methods to test the robustness of model-based indicators for their use as proxies for the state of the food-web to track food-web stability in response to perturbation regimes.

Here, we evaluate the potential of a simulation-based protocol designed to assess the stability of marine food-webs and inform D4, following the GES assessment requirements of the EU-MSFD and its recent updates (European Commission No. 19, 2022). Reliant on principles of ecological stability theory, the protocol provides a framework for the systematic identification of relevant indicators and associated thresholds that can track and signal when perturbation-induced effects in food-web state are most likely to undermine system's stability. Ultimately, this work aims to contribute to the ongoing debate on how trophic models and derived indicators can play a concrete and practical role in the assessment of Descriptor 4 in European seas.

2. Methods

2.1. Modelling approach – An open-ocean and deep-sea case study

We designed a methodological standard (i.e., a simulation-based protocol) to assess the stability of marine food-webs with the Ecopath with Ecosim modelling suite (EwE, Christensen and Walters, 2004) and tested it on an open-ocean and deep-sea system around the Azores (NE Atlantic, Portugal). Ecopath provides a snapshot description of the food-web, as well as the structure and the initial conditions for the Ecosim time-dynamic model (see Christensen and Walters, 2004).

Ecopath is a steady state food-web model that solves a system of algebraic equations. This model is based on the mass-balance of energy flows among biological (functional groups and species) and human components (fishing fleets) of aquatic systems (Polovina, 1984; Christensen and Walters, 2004). Thus, the approach quantitatively represents the structure and functioning of aquatic food-webs (Polovina, 1984). The Ecopath model for the Azores quantifies the trophic interactions between 45 functional groups (FG) for the reference year 1997 (Morato et al., 2016). The model describes the full spectrum of trophic levels (TL) and 11 fishing fleets targeting benthic and pelagic components of the ecosystem. This Ecopath model was used to estimate reference estimates for trophodynamic indicators describing the state of the food-web and to identify biological components that play keystone and dominant roles in the system. Keystone species or functional groups, have an overall high impact in the food-web compared with their relatively low biomass (Paine, 1969; Libralato et al., 2006) and according to model predictions, toothed whales, pelagic sharks, and their main prey (cephalopods, small-size pelagic fish, and bathypelagic fish groups) play such roles in the modelled food-web (Morato et al., 2016). Conversely, the dominant role of some species or groups in a food-web is a result of their relatively high biomass and control on the abundance of other species (McNaughton and Wolf, 1970). Large and small-sized zooplankton groups were predicted to play dominant roles in the modelled food-web.

Ecosim implements a classical dynamic system of differential equations (derived from the Ecopath's system of algebraic equations) where predator–prey interactions are based on flexible functional relationships called foraging arena (Walters et al., 1997; Christensen and Walters, 2004). Each prey group in each predator–prey interaction is divided into biomass fraction readily available (or vulnerable) and unavailable (or invulnerable) for predators, through a vulnerability parameter usually estimated through fitting procedure. The vulnerabilities are key parameters of the foraging arena equations that permit for expressing different functional relationships and have high influence on the dynamic behaviour of the food-web model (Ahrens et al., 2012). The default vulnerability value (=2) implies a mixed mechanism of trophic control (i.e., half population is vulnerable to predator), while low vulnerability indicates a bottom-up control (<2) and high vulnerability (>2) indicates top-down Lotka-Volterra like control (Christensen et al., 2008). Following best practices and guidelines (Heymans et al., 2016), we forced the model with time-series of fishing effort and primary production to represent both human and environmental impacts on the food-webs (e.g., Celić et al., 2018, Chagaris et al., 2020, Bentley et al., 2020). Following, we fitted the model against time-series data from

1997 to 2018, to execute simulations with calibrated vulnerability estimates.

2.2. Conceptual framework adopted to assess food-web stability and inform GES

The simulation-based approach we designed builds on the definition of ecological stability reviewed by Van Meerbeek et al. (2021). According to this, stability is the *overall ability of an ecological system to remain within the same stability domain and to retain its function and structure in the face of perturbations*. Following Carpenter et al. (1999) and Grimm & Wissel (1997), Van Meerbeek et al. (2021) defined the *stability domain* as *the set of state variables and pressure values under which an ecological system returns to a certain reference condition after a perturbation*. Here, the *ecological system* refers to a modelled marine food-web; the *state variables* usually refer to biomasses, but we extended to the trophodynamic indicators that serve as proxies for determining the state of the food-web; the *reference condition* refers to the levels of these indicators in equilibrium and the pressure values are the factors used to simulate the perturbations that might drive significant changes on the state of the food-web.

The above definitions imply an assumption that underlies the approach and is tested in its operationalisation - the simulation of gradients of perturbations lead to effects in trophic structure and functioning of the ecosystem that result in the deviation of state variables from dynamic equilibrium, potentially inhibiting the system's ability to recover. If the modelled food-web is able to withstand the derived effects of the perturbations, the state variables deviate from dynamic equilibrium during the perturbation regime, but return to reference levels once the perturbation is removed from the system. Conversely, if the effects of the perturbation on trophic structure and functioning force the system to exceed the stability boundaries, the state variables do not return to the reference levels, suggesting that the food-web may be in a state that is significantly different from the reference (Fig. 1A). We hypothesise that such stability concepts can be extended beyond state variables and can regard directly the dynamics of indicators that serve as proxies for trophic structure and functioning accordingly: as the indicators gradually deviate from reference levels, in response to a gradient in pressure, the probability of the system recovering to reference levels decreases, while the probability of it not recovering increases (Fig. 1B). This hypothetical behaviour suggests the existence of thresholds indicating the stability boundaries of the food-web, based on probabilities of recovery and non-recovery towards the reference levels. The stability thresholds for the indicators could then be used to delineate a range of values where it is more likely that the food-web will return to the reference state (Fig. 1B). We predict that the range of values will oscillate around the reference levels of the variables in equilibrium. Under the premises of our modelling approach, such a range of values quantitatively describes

marine food-webs at GES.

2.3. A stepwise approach to assess the stability of food-webs

We defined a suite of protocol steps to identify the set of state variables and associated thresholds that allow stability assessments, in line with the assumptions and hypothesis of the approach. The set of state variables should include indicators that describe both trophic structure and functioning (Costanza & Mageau, 1999). Moreover, the trajectories of these indicators, must be robust at tracking perturbation-induced changes in the state of the food-web (i.e., trophic structure and functioning) and at identifying stability boundaries based on threshold values. Indicators with such characteristics can thus be classified as *best proxies for food-web state to track the stability of the system in response to perturbation*. A list of candidate indicators is presented in Table S1. The list includes indicators whose potential to provide useful information for management has been evaluated elsewhere (Heymans et al., 2014, Shin et al., 2010, Samhouri et al., 2009, Link et al., 2015, Libralato et al., 2019, Safi et al., 2019, Fath et al., 2019) and indicators suitable to measure the maturity of ecosystems (Christensen, 1995).

2.3.1. Step 1 – Reference levels for food-web state indicators

As a basis for assessing the stability of the modelled food-web, we defined reference values for each analysed indicator based on the interval of dispersion (1st and 3rd quartile) of values under a dynamic equilibrium regime of five years (Table S2). This regime assumed steady state estimates of total mortality and primary productivity. The range of dispersion of each indicator resulted from a Monte Carlo approach (*Ecosampler*, Steenbeek et al., 2018) that applied random variation to Ecopath input parameters (biomass, production to biomass ratio, consumption to biomass ratio, ecotrophic efficiency, landings and diets), assuming a confidence interval of ± 10 % and a normal distribution. The reference levels for each indicator delimit the standard from which large deviations can be considered significant from a modelling perspective as indicator values are no longer affected by the uncertainty associated to model input parameters. It should be noted that the reference levels represent the best available historical (instead of pristine) state of the food-web in dynamic equilibrium (Morato et al., 2016).

2.3.2. Step 2 - Deviation of food-web from reference levels via perturbation regimes

Subsequently, we forced the food-web in Ecosim to deviate from reference levels by introducing perturbation regimes (press perturbation) over twenty years (Fig. 2). Marine ecosystems and their food-webs can be affected by very different types of perturbation, with potentially different effects on stability. To account for the variability in food-web responses to different types of perturbation, we simulated five perturbation regimes. Regime (R) one to five simulated the impacts of fishing,
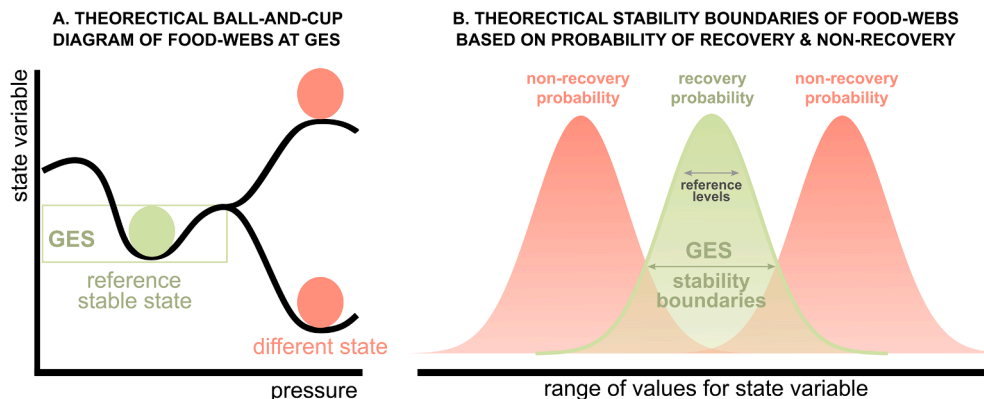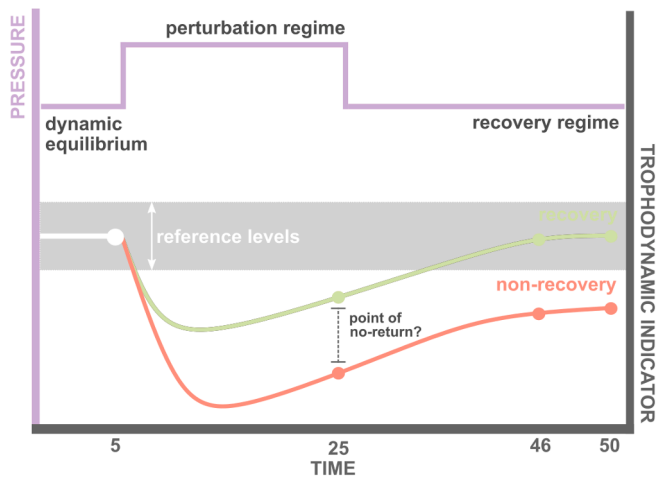

Fig. 1. Schematic representation of the theoretical principles underlying the conceptual framework of the simulation-based protocol to assess the stability of food-webs and inform GES. Here, the *state variable* extends to the indicators that serve as proxies for the state of the food-web.

Fig. 2. Simulation scheme adopted to assess the robustness of indicators and associated thresholds at tracking the stability of modelled food-webs. The system was submitted to regimes of dynamic equilibrium, perturbation and recovery (y-axis, left) so the behaviour of trophodynamic indicators (y-axis, right) could be assessed and thresholds could be identified. The effects of perturbation in food-web state were expected to be reflected in the trajectories of trophodynamic indicators in relation to reference levels in dynamic equilibrium (grey shadow). If the system is suitable to cope with the perturbation, the indicators return to their reference levels, when perturbation is removed (*recovery*). Conversely, if the effects of perturbation undermined stability, indicators are expected to not return to reference levels (*non-recovery*). The threshold value would signal the point of indicator no return to reference levels. The estimates of indicator values were recorded in each model run for year 5 (dynamic equilibrium), year 25 (end of perturbation regime) and years 46–50 (median value at the end of the recovery regime).

low primary productivity, fishing and low primary productivity, impacts on keystone groups and impacts on dominant groups, respectively, on food-web state. To simulate fishing effects (R1), we increased the levels of fishing effort of all modelled fleets, relatively to reference estimates, while assuming reference estimates of primary productivity. To simulate

effects on productivity (R2), we decreased the production to biomass ratio (P/B) of primary producers (phytoplankton and algae), assuming reference levels of fishing effort. To simulate the combined effect of fishing and low primary productivity (R3) we increased fishing effort along with a decreases of P/B ratios of primary producers. To simulate the effects on structuring species, we increased the mortality of modelled keystone (R4) and dominant groups (R5). R2, R4 and R5 might capture fishing effects derived from reference effort levels. We simulated the effects of each perturbation regime under five magnitudes of pressure, ascending from very low to very high pressure (Table S3). Magnitudes specific for different pressures resulted from trial exercises (not shown here), and, therefore, are specific for the case-study. The same Monte Carlo approach used to delineate the reference levels of each indicator was used to generate replicates of each indicator in each perturbation regime. We considered a hundred replicates per simulation a good trade-off between statistical significance of results and processing time. Therefore, simulations involved a total number of 2500 trials (5 perturbation regimes x 5 magnitudes of pressure x 100 replicates of estimates).

2.3.3. Step 3 – Systematic evaluation of indicators for stability assessments

We systematically evaluated the trajectories of trophodynamic indicators in the five perturbation regimes to identify those that best track changes in trophic structure and function caused by perturbation. This process was based on a scoring scheme that ranked the performance of candidate indicators meeting quantitative targets (Fig. 3). The targets ensured that indicator behaviour was responsive (criterion A), consistent (criterion B) and reliable (criterion C) in informing perturbation-derived changes in food-web state. The rationale for the criteria and the respective targets fit some of the standards that have been introduced to identify practical, ready-to-use food-web indicators (Tam et al., 2017). Responsiveness here refers to the ability of the indicator to be sensitive to the gradient in the magnitude of pressure and to deviate from reference levels regardless of the driver or mechanism of perturbation (e.g. top-down or bottom-up impacts). Consistency refers to the ability of the indicator to show clear trajectories in response to perturbations that allow trend monitoring. Specifically, the indicator should

Fig. 3. Criteria adopted in the systematic evaluation of candidate indicators to select best proxies of food-web state for the stability assessment and respective description of quantitative targets. The performance of each indicator was assessed based on median value of indicator's predicted distribution (from indicator's replicates), at the end of the perturbation regime. The maximum score possible to be reached by each indicator was 75 points (25 attributed to each criterion, considering 5 regimes of perturbation x 5 magnitudes of pressure). Examples illustrate a hypothetical behaviour of an indicator that would perfectly meet the quantitative targets.

show unidirectional trends in response to a driver of perturbation and provide a clear sense of the magnitude of change (i.e., higher pressure intensity amplifies the indicator's response compared to reference levels). Accordingly, the trend of the indicator relative to reference levels could be used as a warning signal of approaching food-web stability boundaries. Finally, the reliability of the indicator refers to the influence of model quality on the trajectory of the indicator. Ideally, the trajectory of the indicator should be consistent across alternative uncertainties in the data collection (i.e. low or high levels of uncertainty) to ensure the accuracy of the information they provide about the direction of changes in the state of the food-web. Indicators obtaining the highest performance rankings were classified as best to track perturbation-induced changes in food-web state (hereafter also named best state indicators).

2.3.4. Step 4 – Classification of food-web stability condition after perturbation regimes

After perturbation, we simulated a recovery regime by re-introducing initial pressures in the system (i.e., steady-state estimates for mortality and productivity) over a period of twenty-five years (Fig. 2). This simulation scheme allowed the food-web to (eventually) reach dynamic equilibrium and furthermore assess the ability of food-web state in recovering from the effects of perturbations. We assumed twenty-five years as the upper limit for the recovery time because perturbation-induced effects that require a recovery period of five to twenty years can be considered more than temporary (FAO, 2009). We assigned a classification of stability condition to each replicate of the trophodynamic indicators that serve as best proxies of food-web state to track perturbation-induced changes. The classification established the ability of each replicate in returning to reference levels (at the end of the recovery regime). First, we calculated the median value of each indicator replicate, over the last five years of the recovery regime. Following, we recorded the values obtained at the end of the perturbation regime (year 25) and assigned the classification of *recovery* to replicates which median value fell within reference levels and *non-recovery* to replicates which median value fell beyond reference levels. We excluded from the analysis Monte Carlo replicates whose indicator estimates did not initiate within reference levels.

2.3.5. Step 5 – Validation of the probability approach to detect food-web stability boundaries

We analysed the distribution of Monte Carlo replicates of best state indicators that recovered and did not recover from the perturbation regime and estimated the probabilities of return towards reference levels in dynamic equilibrium (Supplementary Material). First, we estimated the probability of the system *recovery* or *non-recovery* from perturbation when the values of the indicators (under perturbation) oscillated within and beyond reference levels. We then estimated the same probabilities for each level of the simulated gradient of magnitude of pressure (from very low to very high pressure). In order for the best food-web state indicators to provide information on the probability of *recovery* or *non-recovery* from simulated perturbations, we expected i) the probability of *recovery* to be higher than the probability of *non-recovery* when the indicator values are within the reference levels; (ii) that the probability of *non-recovery* is higher than the probability of *recovery* when indicator values are beyond the reference levels; (iii) that the probability of *recovery* progressively decreases in response to the gradient of pressure; and (iv) that the probability of *non-recovery* progressively increases in response to the gradient of pressure (Figure S2, left panel). We estimated probabilities considering the frequency of an event - i.e., the number of times an indicator recovered from perturbation when the system was under high magnitude of pressure divided by the total number of simulation trials.

2.3.6. Step 6 – Evaluation of the robustness of indicators in detecting the stability boundaries of the food-web

We estimated stability thresholds for the best state indicators based on signal detection theory (SDT) methods (Murtaugh, 1996, Rice, 2003) by comparing the ability of each indicator to correctly detect *non-recovery* conditions (sensitivity, true positive) against the ability to correctly detect *recovery* conditions (specificity, true negative), at the end of the perturbation regime. This assessment was based on the test statistics of the Receiver Operating Characteristic (ROC) Plot. The thresholds for each indicator corresponded to the optimal cut-off point that maximises the difference between true and false predictions across all possible thresholds. This was assessed considering the maximum value of the Youden index (Youden, 1950). If the indicator is sensitive in diagnosing *recovery* and *non-recovery* conditions, a Youden index value above 0.5 should be expected. We used the area under the ROC curve (AUC) as a measure of the indicator's performance in identifying both conditions of the food-web. If the indicator perfectly dichotomises the two conditions of *recovery* and *non-recovery,* an AUC performance statistic of 1 is expected. In ecological studies, AUC values ≥ 0.8 indicate an excellent indicator response and ≥ 0.7 an acceptable response (Hale and Heltshe, 2008). An AUC value of 0.5 indicates predictive discrimination no better than random guessing (Elith et al., 2006). Statistical analysis was performed using the R package *cutpointr* (Thiele & Hirschfeld, 2021). The strengths and limitations of SDT methods are well documented and can be found in Macmillan & Creelman, 2004.

We estimated two additional metrics to measure the predictive power of indicators, taking into account the prevalence of *non-recovery* condition in the analysed population of replicates. These were the positive predictive value (PPV: the probability of true positives) and the negative predictive value (NPV: the probability of true negatives) (Murtaugh, 1996). Low prevalence means that the proportion of the population of replicates presenting a *non-recovery* condition is small. In this case, the PPV is lower and the NPV is higher. Since the PPV is related to the specificity of the indicator and the NPV to the sensitivity, there are trade-offs between the two measures, namely around the risk of the indicator providing false information. We analysed the risks associated with each indicator by quantifying the probability of the indicator providing a false positive (low specificity) and a false negative (low sensitivity) (Olliaro & Torreele, 2021). These were expressed numerically as 1-PPV ("false detection rate", FDR) and 1-NPV ("false omission rate", FOR) respectively. Ideally, the indicator should have low rates of false detention and omission (i.e. < 10 %). However, in the context of the present study, the risks associated with false omission of the non-recovery state are more harmful than the risks associated with false detentions. In summary, if the indicator is sufficiently robust in detecting different stability conditions of the food-web based on threshold values, the following quantitative targets should be expected: AUC value ≥ 0.8, Youden index > 0.5, FOR and FDR < 10 %. ROC curves, associated statistics and predictive probabilities were estimated for the lower and upper bounds of the best-scoring indicators based on how the perturbations affect the trajectory of the indicator relative to the reference estimates (i.e. decreasing trend: lower bound, increasing trend: upper bound).

2.4. Applications of the simulation-based approach

We used the best food-web state indicators and respective stability thresholds to identify i) the impact of loss of stability on the balance between trophic levels and ii) the biological components of the food-web that are most affected in scenarios of stability loss. Scenarios of stability loss correspond to Monte Carlo trials where indicators' values were found beyond the stability threshold (*indicator collapse*, for simplicity). To quantify the impact of stability loss on the balance between trophic levels, we compared the proportions of total biomass per trophic level (as derived from the trophic spectra, Libralato and Solidoro, 2010) when the indicator values were within the reference levels in dynamic

equilibrium and beyond the stability threshold. To identify the functional groups most affected by the loss of stability, we estimated the probability (i.e., frequency) with which the biomass of each modelled functional group returned and did not return to the respective reference levels, for those scenarios. Reference levels of biomass for each functional group were estimated as described in Section 2.4.1 (Table S4).

3. Results

The application of the simulation-based protocol to assess the stability of the open-ocean and deep-sea food-web provided two important results: the identification of the most robust indicators to track perturbation-induced changes in the state of the food-web and the detection of stability boundaries based on threshold values on such indicators. Furthermore, the identified indicators and thresholds were used to highlight the impact of loss of stability on the balance between trophic levels and the most affected biological components in such scenarios.

3.1. Best indicators to track perturbation-derived changes in food-web state

To identify indicators that best track perturbation-induced changes in trophic structure and functioning, we evaluated their performance in meeting certain behavioural criteria, under perturbation regimes. For an indicator to exhibit responsive behaviour to perturbation (criterion A), we expected the median value of the indicator's predicted distribution to deviate from reference values only above a moderate magnitude of pressure (Fig. 3, A.1). Indicators that quantify transfer efficiency (TE) through the food-web and measure the average trophic level of the community (TL Community) performed well at this level, namely when fisheries impacts were simulated with (R3) and without (R1) variation in primary production (Table 1 and Fig. 4). However, impacts on the

structuring groups led to deviations of the indicator values from the reference levels (R4 and R5) (Fig. 4), at very low and low magnitude of pressure (Fig. 4), suggesting high sensitivity of indicators to such perturbation regimes. To fully meet the responsiveness criterion, we also expected the indicators to respond to all simulated perturbation regimes at a certain magnitude of pressure (Fig. 3, A.2). This behaviour was only confirmed for the total transfer efficiency (TE) indicator estimated as an emerging property of the trophic spectra (Total TE-Spectra) and the parameters describing the cumulative biomass curve (indicators for steepness and inflection point (Table 1 and Fig. 4). Conversely, the TE indicators (estimated on the basis of flows of matter) and the TL Community indicator did not respond to fisheries impacts when simulated as a sole driver of perturbation and to impacts on keystone groups, respectively.

For indicators to reliably track and inform state changes, we expected the trajectories to be consistent (criterion B) in the perturbation regimes. In particular, we expected indicator values to consistently increase or decrease in response to a specific perturbation regime, compared to the reference levels in dynamic equilibrium (Fig. 3, B.1). The indicators based on TL Community and TE achieved the target perfectly, with the exception of the Total TE-Spectra indicator (Table 1 and Fig. 4). In this case, the trajectory of the indicator at very low and low fishing-induced pressure differed from that estimated at high and very high pressure (Fig. 4). However, there was a consistent relationship between the pressure level and the deviation from the reference levels. This means that higher pressure in the system results in a larger deviation of the indicator from reference levels, which grounds consistency to the indicator (Table 1, B.2).

Finally, we assessed the consistency of indicator trajectories under different assumptions regarding uncertainty in the sampling of model input data (criterion C) (Fig. 3). When assuming low sampling uncertainty (using high-quality data for parameterizing the model), the simulated perturbations caused the Total TE-Spectra indicator values to

Table 1
Performance of candidate indicators meeting the quantitative targets for each criterion that establishes the robustness of the indicator in tracking impacts on food web status caused by perturbation. The final score is expressed as a % of success in achieving the maximum score.

Ranking	Indicator Name	A1. Responsiveness	A2. Responsiveness	B1. Consistency	B2. Consistency	C. Reliability	Final Score (%)
1	Total Transfer Efficiency (flows)	9.5	10	12.5	10	25	89.3
2	PP Transfer Efficiency (flows)	9	10	12.5	10	25	88.7
3	Total Transfer Efficiency (Spectra)	10.5	12.5	5	12.5	25	87.3
4	Detritus Transfer Efficiency (flows)	8.5	10	12.5	7.5	25	84.7
5	Trophic Level Community	8.5	10	12.5	7.5	25	84.7
6	Biomass at Inflection Point of cumBiom curve	8	10	12.5	5	25	80.7
7	Average Mutual Information	8.5	7.5	7.5	7.5	25	74.7
8	Total System Biomass	8	7.5	7.5	7.5	25	74.0
9	Development Capacity	9.5	7.5	5	7.5	25	72.7
10	Ascendency	8.5	7.5	5	7.5	25	71.3
11	Entropy	8	7.5	5	7.5	25	70.7
12	Respiration	8	7.5	5	7.5	25	70.7
13	Total Biomass to Throughput ratio	7.5	7.5	5	7.5	25	70.0
14	Average Path Length	7.5	7.5	5	7.5	25	70.0
15	Finn Cycling Index	7	7.5	5	7.5	25	69.3
16	Detritivory to Herbivory ratio	8.5	7.5	7.5	7.5	20	68.0
17	Overhead	7.5	7.5	5	5	25	66.7
18	Primary Production	7.5	5	7.5	5	25	66.7
19	Relative Ascendency	7	5	7.5	5	25	66.0
20	Relative Overhead	7	5	7.5	5	25	66.0
21	Redundancy	8	7.5	5	5	20	60.7
22	Steepness of cumBiom curve	10	12.5	2.5	5	15	60.0
23	Inflection Point of cumBiom curve	9.5	12.5	2.5	5	15	59.3
24	Flows to Detritus	6.5	5	7.5	0	25	58.7
25	Total System Throughput	6.5	5	5	2.5	25	58.7
26	Total System Production	7.5	5	5	5	20	56.7
27	Predatory Cycling Index	7	7.5	7.5	5	15	56.0
28	Robustness	7	5	7.5	5	15	52.7
29	Total Primary Production to Total Respiration ratio	6	2.5	5	0	20	44.7
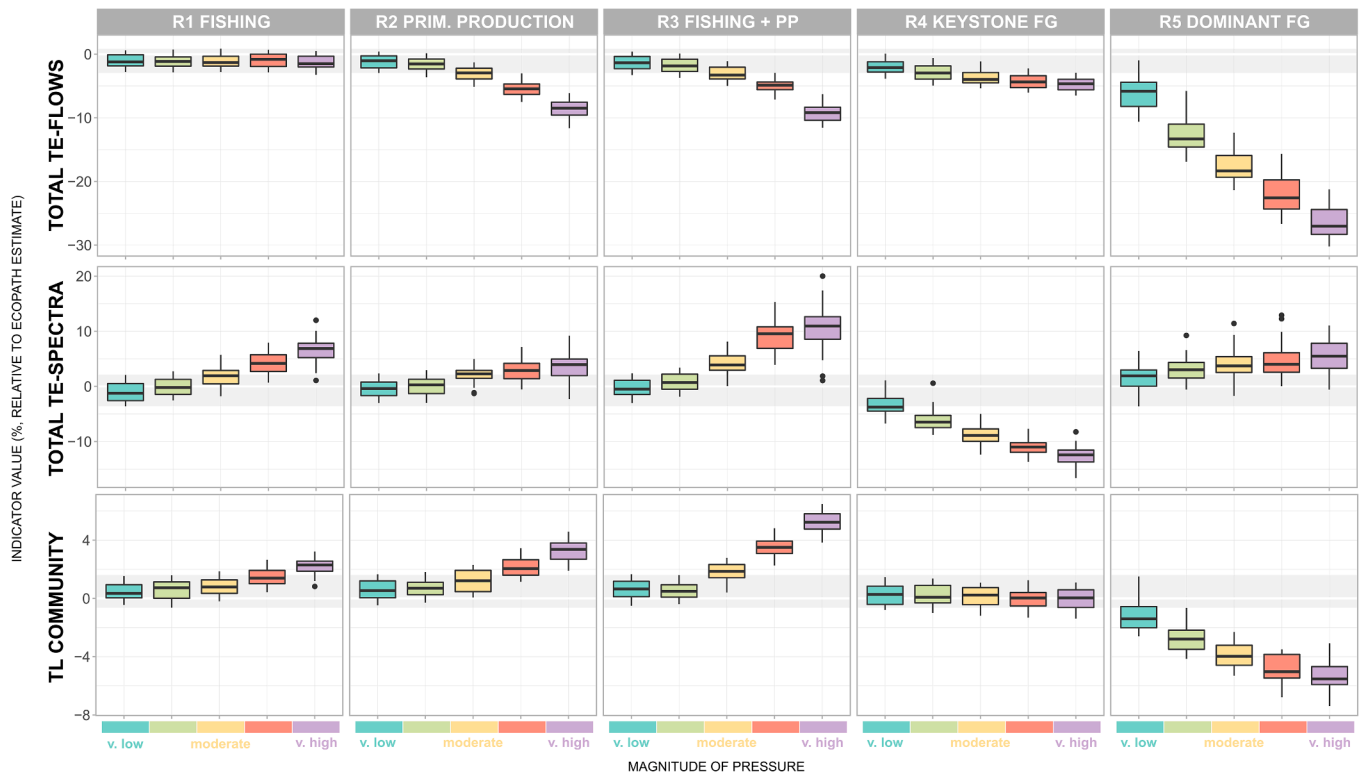
Fig. 4. Boxplots showing the relative response of best state indicators to the full set of simulated perturbation regimes. Results illustrate indicator responses (y-axis) to a gradient of magnitude of pressure (x-axis), at the end of the perturbation regime. The white line highlights the reference estimate of each indicator (0%; from Ecopath model), while the grey shadow demarks the range of values that delineate the reference levels of the indicator, based on dispersion in dynamic equilibrium. Model predictions assume low sampling uncertainty in input data.

increase compared to the reference estimate (Ecopath), except when the perturbation effects resulted from increased mortality of keystone groups (Fig. 4). A similar pattern was observed for the TL Community indicator, where the decreasing trend was caused by increased mortality of dominant groups. This suggests that both indicators have upper and lower bounds (Fig. 4). In contrast, TE indicators based on matter flows tended to decrease during perturbation regimes, indicating the presence of a single lower limit (Fig. 4). These indicator trajectories, relative to reference (Ecopath estimates), were also observed under the assumption of high sampling uncertainty (low-quality model). This result suggests that the information they provide about the food-web's state remains unaffected by model quality and can be considered reliable (Fig. 5).

The overall behaviour of TE-based indicators and the TL Community indicator was translated into a capacity to meet > 84 % of the quantitative targets set for the behavioural criteria (Table 1). These results suggest that they are indeed the most robust in tracking perturbation-induced changes in food-web state, and hence they were selected for threshold estimation. For simplicity, we highlight the results for the Total TE-Flows, Total TE-Spectra and TL Community indicators, as the responses based on the TE from PP and TE from Detritus indicators were very similar to those predicted for the Total TE-Flows indicator (Table 1).

3.2. Robustness of best state indicators at detecting stability boundaries

Before estimating the thresholds on best state indicators, we analysed the distribution of Monte Carlo replicates that recovered and did not recover from the perturbation to validate the assumptions
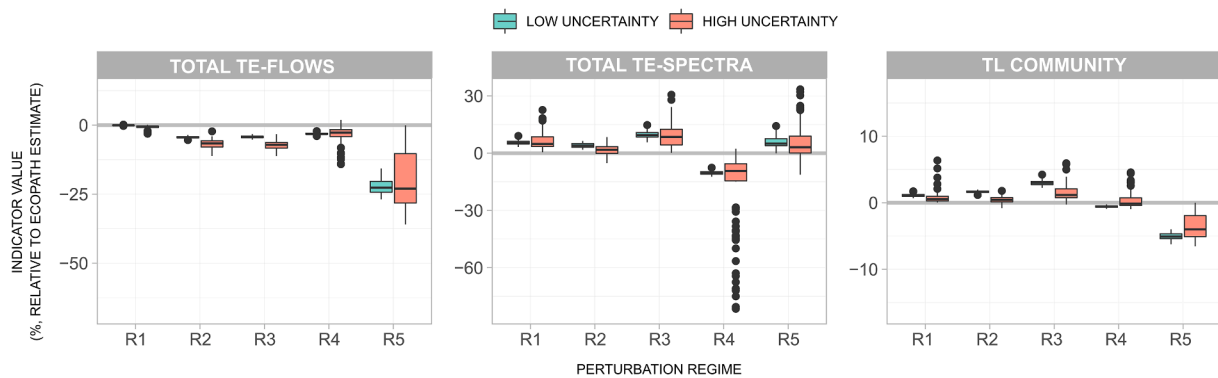

Fig. 5. Boxplots showing the relative response of best state indicators to the full set of simulated perturbation regimes (x-axis) when the sampling uncertainty associated to model input data is low (high quality model) and high (low quality model). Results illustrate indicator responses (y-axis) to a high magnitude of pressure, at the end of the perturbation regime. The grey line highlights the Ecopath estimate of each indicator (0%).

underlying the approach adopted to detect the stability boundaries of the food-web (Supplementary Material). Following, we estimated stability thresholds for the best state variables based on the Youden index. Accordingly, the threshold that best distinguished between recovery and non-recovery conditions of the modelled food-web based on the Total TE-Flows indicator was −9.17 % (Fig. 6). The thresholds for the Total TE-Spectra indicator were −6.93 % and 5.29 %, corresponding to the lower and upper bounds of the indicator, respectively (Fig. 6). The thresholds for the TL Community indicator were −4.86 % and 2.08 %, corresponding to the lower and upper bounds of the indicator, respectively (Fig. 6). The estimated Youden index was 0.59 %, 0.52 % and 0.55 % for the total TE -Total Flows, TE -Total Spectra and TL

Community (upper bound) indicators, respectively. These values indicate a low occurrence of incorrect results (i.e. false negatives or false positives). In contrast, the lower bound of the Total TE -Spectra and TL-Community indicators were found to have a low Youden index (0.08 % and 0.32 %, respectively), suggesting low confidence in the estimates. As expected, the thresholds were found beyond the reference values of each indicator in dynamic equilibrium. Both the reference levels and the thresholds are the percentage change from the Ecopath estimate for each indicator.

The SDT analysis underlined the robustness of the Total TE-Flows indicator to delineate the stability boundaries of the food-web in response to perturbation (Fig. 7). The indicator showed excellent results
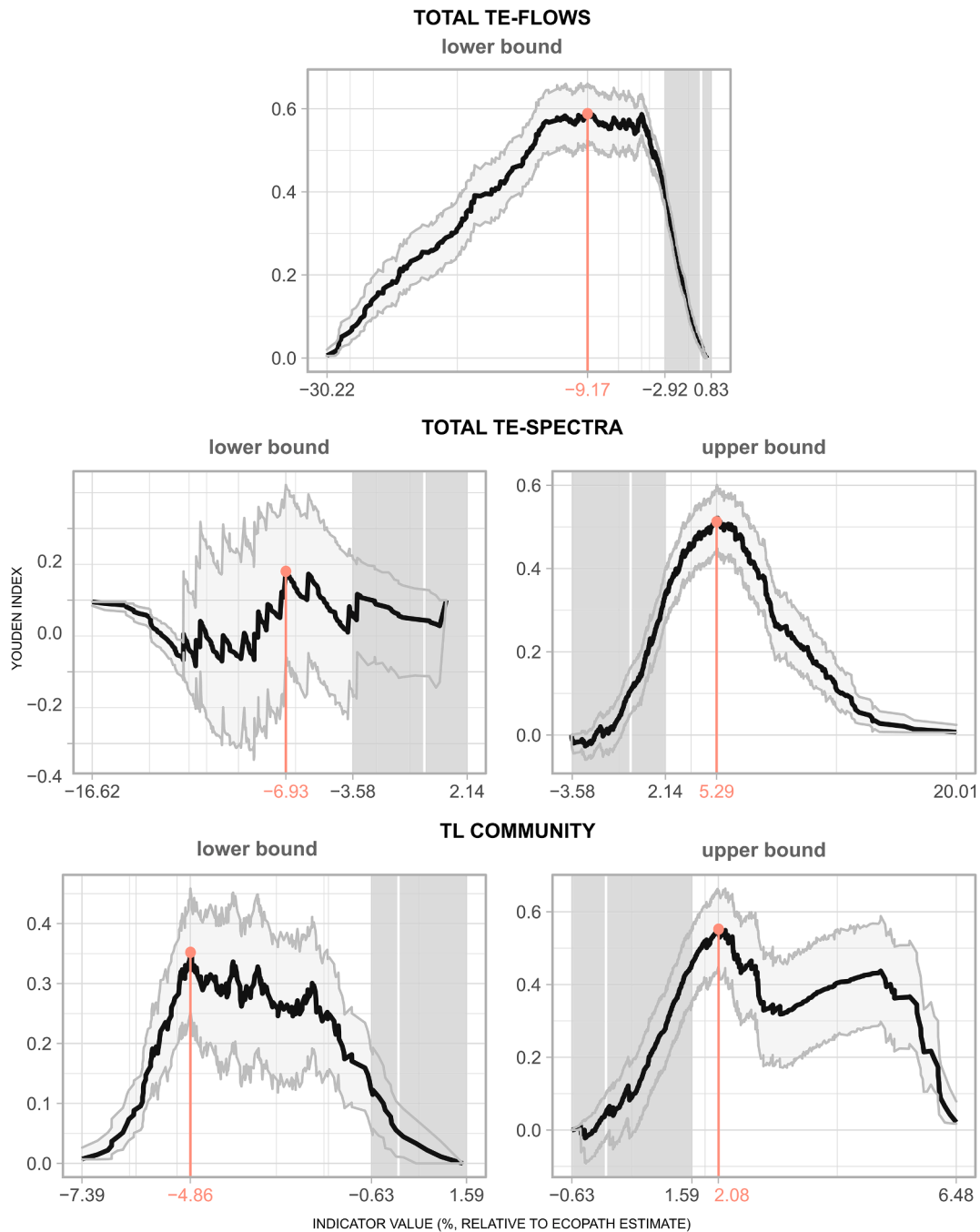
Fig. 6. Identification of stability threshold values (in red) for the best proxies of food-web state (x-axis) based on indicator ability to distinguish stability conditions of the food-web (y-axis, Youden Index). The white line highlights the reference estimate of each indicator (0%; from Ecopath model), while the grey shadow demarks the range of values that delineate the reference levels of the indicator, based on dispersion in dynamic equilibrium. The variability around the Youden Index estimates for each indicator derives from bootstrapping methods (Thiele & Hirschfeld, 2021).
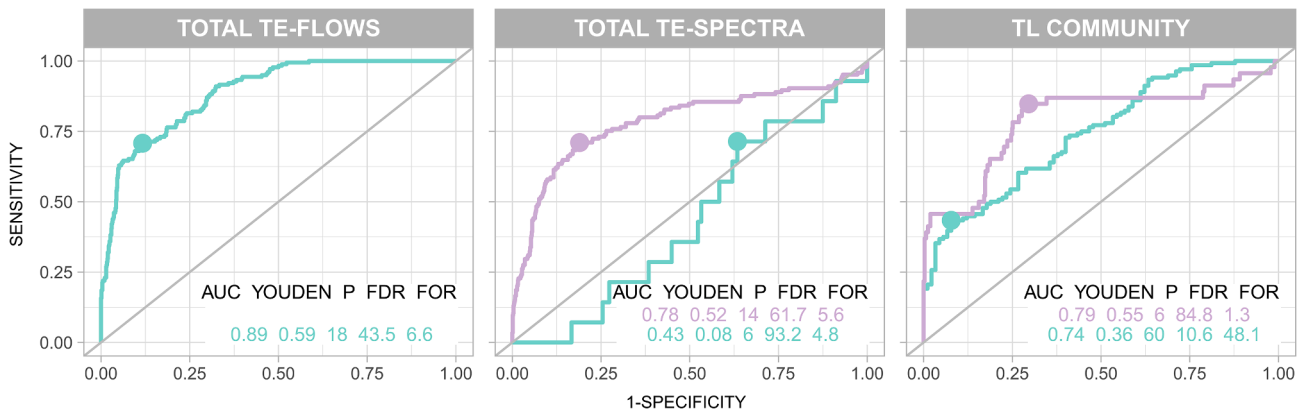
Fig. 7. ROC (receiver operating characteristic) curves, associated statistics (AUC and Youden Index) and predictive probabilities (P, prevalence; FDR, false detention rate; FOR, false omission rate) obtained for the best proxies of food-web state while discriminating *recovery* and *non-recovery* conditions of the food-web. The lower and upper bounds of each indicator are represented in green and purple, respectively. The most accurate thresholds for distinguishing stability conditions, based on Youden Index, are denoted with a dot.

(AUC = > 0.89) while discriminating between recovery and non-recovery stability conditions (Fig. 7). At its respective prevalence value (18 %), the indicator was accurate in avoiding false-negative results for non-recovery conditions (FOR 6.6 %), but the associated probability of falsely predicting non-recovery conditions (FDR) was estimated to be over 25 %. The results based on the summary metrics for the diagnostic accuracy of the Total TE-Spectra and TL Community suggest a lower but still acceptable robustness of these indicators in identifying stability boundaries when focusing on their upper limits. However, the ability of Total TE -Spectra to distinguish stability conditions when the indicator presents a decreasing trend compared to reference levels is no better than random estimation. The negative values of the TL Community indicator seem to be able to correctly distinguish both stability conditions (AUC 0.74), but the associated probability of missing non-recovery conditions is high (FOR 48.1 %). Assuming the most robust thresholds, the range of values at which the food-web is more likely to return to the steady state of reference corresponds to [-9.17 %, 0.83 %], [-3.58 %, 5.29 %] and [-0.63 %, 2.08 %] for the indicators Total TE-Flows, Total TE-Spectra and TL Community, respectively (Fig. 7).

3.3. Effects of stability loss on the balance between trophic levels

To determine the effects of the loss of food-web stability on the balance between trophic levels, we compared the trophic spectra estimated in scenarios where state indicators have collapsed and in dynamic equilibrium. We found that the loss of food-web stability can have important effects on the balance between trophic levels. These effects can manifest themselves in two different ways depending on the regime of perturbation (Fig. 8). Fishing-induced perturbations (R1-R3) caused an increase in biomass at trophic level 3 and higher, while biomass at lower trophic levels decreased (Fig. 8, left panel). However, the overall distribution of biomass across trophic levels remained pyramid-shaped, similar to the shape estimated in reference conditions. On the other hand, perturbations that affected dominant groups led to some restructuring of the food-web's shape (Fig. 8, right panel). However, when stability was lost, a higher proportion of total biomass was concentrated at trophic levels 3–3.5 compared to the level below. Consequently, the proportion of biomass at trophic level 2.0 and above 3.5 was predicted to decrease and increase, respectively, compared to the reference values.
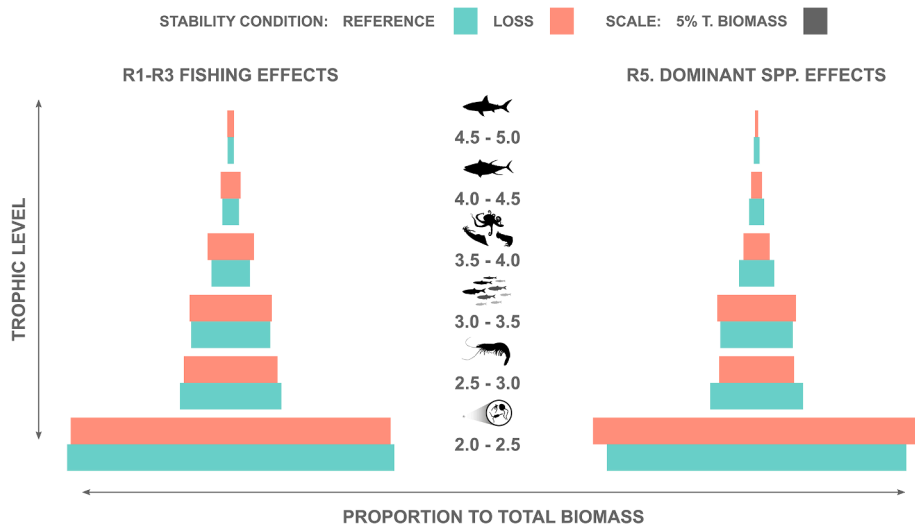
Fig. 8. Representation of estimated proportion of total biomass per trophic guild (as derived from trophic spectra) under reference conditions (green bars) and scenarios of stability loss (red bars) driven by fishing and impacts on species that play dominant roles in the food-web. Presented proportions of biomass correspond to averaged estimates among the full set of trials where indicator values were found beyond the stability thresholds. Illustrations from Valter Medeiros.

3.4. Most affected biological components in scenarios of stability loss

Our simulations show that pelagic (e.g., tunas, pelagic sharks) and benthic predators (e.g. large demersal fish group, *Phycis phycis, Beryx decadactylus, Conger conger*) are among the most likely biological components to not return to the level of the reference biomass when the indicators Total TE-Flows, Total-TE Spectra and TL Community exceed their respective stability thresholds (Fig. 9). In such cases, perturbations caused the biomass of these components to decrease compared to the Ecopath reference estimate. However, we also found that perturbation led to significant increases in the biomass of certain groups occupying lower to middle trophic levels in the food-web. This was particularly true for medium-sized pelagic fish, demersal fish and shallow-water fish, whose biomass was above their reference levels when the indicators based on TE collapsed (Figure S3). These findings suggest that perturbations are expected to have a negative impact on high trophic levels (due to low turnover rates) but a positive effect on lower trophic levels.

4. Discussion

4.1. Strength and limitations of the approach

Stability assessments of modelled marine food-webs could be a concrete and practical way to support the monitoring of Descriptor 4 in European Seas, especially because the scientific community continues to develop high-resolution trophic models of great potential to be used operationally. In this context, the evaluation of the simulation-based protocol we designed and applied in this study suggests it might be a useful standard approach to systematically identify the indicators and associated thresholds with the greatest potential to perform such assessments.

Our simulations clearly showed variability in the performance of model-based indicators at tracking and signalling when perturbation-induced effects in food-web state are most likely to undermine system's stability. Among the analysed indicators, transfer efficiency (Total TE-Flows and Total TE-Spectra indicators) and average trophic level of the community (TL Community) appear to be the best proxies for trophic functioning and structure, respectively, to perform stability assessments. The validity of this conclusion is supported by the demonstrated capacity of indicators in meeting quantitative criteria that ground robustness at i) exhibiting a sensitive response to a gradient of pressure and to various perturbation mechanisms, ii) consistently reflecting

trajectories that indicate the direction of change in the food-web state, and iii) distinguishing disparate stability food-web conditions based on threshold values. Furthermore, these indicators possess a strong theoretical foundation, which ensures appropriate utilization in ecosystem-based management approaches (e.g., Eddy et al., 2021; Fath et al., 2019, Safi et al., 2019). These results highlight the importance of adopting quantitative approaches to ensure objectivity, reproducibility and precision in the selection of relevant indicators for tracking specific attributes or properties of marine food-webs (Otto et al., 2018, Rossberg et al., 2017, Samhouri et al., 2009).

When the modelled food-web was subject to regimes of perturbation, the mentioned indicators tended to deviate from reference estimates, and the values they assumed provided information on the severity of the impacts. Overall, higher pressure in the system led to greater deviation from the dynamic equilibrium estimates, resulting in a lower probability of system recovery and a higher probability of system non-recovery towards the state of reference. Thus, the indicator distance from the reference levels is indeed associated with distinct probabilities of system *recovery* and *non-recovery*, which permits to validate the assumptions underlying the stability approach. Moreover, this behaviour allowed identifying threshold values in the indicators that distinguish system *recovery* and *non-recovery* and therefore demark the point of no-return towards the state of reference of the food-web.

The identification of stability thresholds is an ecologically significant outcome because it allows us to determine a level at which the indicators signal a warning how the state of the food-web is affected by perturbations. Specifically, when this level is exceeded, the indicator warns that perturbations are driving such profound effects on trophic structure and functioning and that system's stability has been compromised. Based on this result, a scale can be established to measure the impact of perturbations on the state of the food-web, in relation to the reference state: insignificant impacts when indicator values fluctuate within the range of the reference levels; significant impacts when values fall between the reference levels and the stability threshold; and highly significant impacts when values exceed the stability threshold. In the absence of other specific quantitative targets describing food-webs in GES, the detection of indicator trends approaching estimated stability thresholds could serve as the best available warning for *adverse effects*, manifested in the form of deviations from stable boundaries of the food-web.

The described scale of ecological significance of impacts on the food-web was found most reliable when based on the lower (and single)
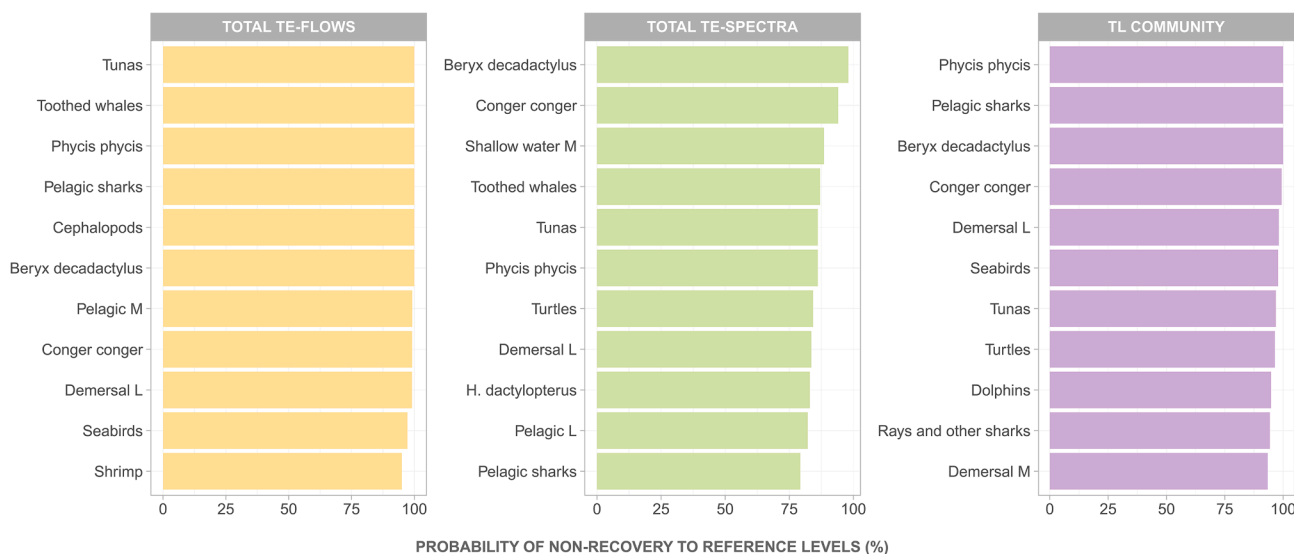

Fig. 9. Probability (i.e., frequency) of non-recovery to reference levels of biomass estimated per modelled biological component in scenarios of stability loss. S, M and L stand for small, medium and large-size, respectively.

bound of the Total TE-flows indicator. Estimated projections of this indicator matched most of the properties that the state indicators should have under the premises of our stability assessment approach. In fact, this was the indicator that most closely matched the expected patterns of the probability of *recovery* and *non-recovery* of the state of the food-web in response to the gradient in the magnitude of the pressure (Fig. 7). This behaviour explains the good performance of the indicator in meeting the targets for the ROC statistics (AUC and Youden index) and the estimated low probability of omitting *non-recovery* conditions of the food-web. The results based on the same statistics show that thresholds can also be estimated for the upper bounds of the Total TE-Spectra and TL Community indicators. However, their respective lower bounds have limited applicability, as they were unable to correctly distinguish the state of the food-web (Total TE-Spectra indicator) and are highly likely of omitting *non-recovery* conditions of the system (TL Community indicator, FOR 48 %).

Despite its strengths, the Total TE-Flows indicator also presented limitations, centred on the inability to respond to fishing-induced effects, when simulated as the sole driver of perturbation. Indicators based on flows of matter are particularly sensitive to perturbations that grasp effects at the level of total system biomass (Saint-Béat et al., 2015 and references therein). As typical from open-ocean food-webs, the greatest proportion of modelled system biomass is concentrated at the bottom of the food-web (60 % of total biomass at trophic level 2.38) (Link et al., 2015), implying that pronounced changes in total system biomass are more likely triggered by perturbations on low trophic levels (e.g., changes in primary production). Therefore, for the indicator to detect top-down impacts, predator-release must trigger cascading effects down the food-web that are strong enough to cause pronounced changes in the total biomass of the system. The fact that the indicator responds to impacts on keystone species (considered top-down effects in our model) but not intense fishing, suggests that the signature of trophic cascades in our model might be dampened when the pressure is distributed over a wide spectrum of trophic levels such as those exploited by fishing (Andersen and Pedersen, 2010). This means that the consequences of the cascading effects caused by fishing on the stability of the food-web can best be diagnosed by analysing the trajectories and the crossing of the thresholds of the Total TE-Spectra and TL Community indicators.

When performing such diagnostics, it is important to take into account the characteristics and limitations inherent to the stability approach here evaluated. Crossing the thresholds warns that a greater restoration effort is needed for the modelled food-web to return to the dynamic equilibrium of reference, assuming a specific i) length of perturbation exposure, ii) gradient of magnitude of pressure, iii) type of perturbation (press), iv) set of drivers of perturbation, v) reference state and recovery condition and vi) deterministic modelling approach. The approach is replicable, objective and both theoretically and statistically sound, although the robustness of the estimated thresholds could be further evaluated by using alternative perturbation and recovery lengths, pressure intensity gradients, drivers and types of perturbation (e.g., pulse vs continuous; gradually increasing vs abrupt change). Furthermore, although multiple simulations and model parameterisations allowed accounting for uncertainty and variability in results, the approach does not account for all the factors that affect real ecosystem dynamics and might limit models' ability to predict future ecological states (e.g., stochasticity) (Planque, 2016 and references therein; Storch et al., 2017). It is also important to note that both architecture of the food-web network (i.e., aggregation of functional groups and species in the model and their life history traits) (Pinnegar et al., 2005) and Ecosim fitting approach affect recovery patterns of the modelled system. Yet, quantifications of the extent to which such features influence model projections are difficult to obtain, particularly in the absence of ensemble modelling frameworks, performing similar sets of simulations with alternative modelling approaches and network configurations.

4.2. Operational applications of the approach to inform Descriptor 4

Having identified the most robust state indicators of the food-web and associated stability thresholds, we explored a number of applications that might be useful to extend the knowledge of food-web dynamics derived from trophic models. One immediate application of the approach is to build scenarios on the mechanisms underlying changes in the state of the food-web, based on the direction of change in the indicator trajectory in response to perturbation. For instance, scenarios directly involving fishing effects (S1 and S3) resulted in increasing trends on the Total TE-Spectra and TL Community indicators and decreasing trends on the Total TE-Flows indicators. The contrasting trends of the TE-based indicators (which is explained by their calculation methods) provide complementary information on energetic changes within the food-web. The decreasing trend in the Total TE-Flows indicator suggests that perturbations may reduce the amount of energy reaching higher trophic levels, which is consistent with findings from other studies (Andersen et al., 2009, Anh et al., 2015, Heymans, 2003). Conversely, the increasing trend in the Total TE-Spectra indicator suggests that energy flows might become more restricted and carried out by species with shorter life expectancy but efficient at transferring energy through trophic levels (Du Pontavice et al., 2020, Maureaud et al., 2017). The TL Community indicator suggests that species and functional groups occupying middle trophic levels (such as TL 3–3.5, including small demersal and bathydemersal fish and medium-sized shallow water, demersal and pelagic fish) may play such novel roles in energy transfer.

The increasing trend of the TL Community indicator in certain regimes is explained by mixed contributions of bottom-up and top-down effects induced by fishing pressure, which is not evenly distributed across trophic levels in the food-web, as mentioned before. In the modelled system, the pelagic fleets target species of a wide range of trophic levels (e.g., swordfish, tunas, jacks, small-sized pelagic fish), while deep-sea fishing is more specific and mostly targets trophic levels $>= 4$ (e.g., conger eel, seabreams, black scabbardfish) (Morato et al., 2016). If fishing effort is pressuring pelagic species in multiple TL and demersal species in a specific part of the food-web, the impacts can be especially strong for benthic groups. In this case, when fishing pressure is leading the decline of high TL benthic species, predation release favours the biomass of prey of TL between 3 and 3.5. In contrast, the scenario involving impacts on dominant species (R5) resulted in a decreasing trend of the indicator, that can be expected when the perturbation produces top-down cascading effects in the system (e.g., Branch et al., 2010). We recall that the described food-web responses to perturbation are hypothetical and can only be validated by comparison with empirical data. Further work could be carried out to compare the results with possible empirical "natural experiments" or at least to carry out analysis across temporal and spatial scales. Notably, these indicators can be calculated directly over time-series data, albeit after appropriate transformations and assumptions (Libralato and Solidoro, 2010).

Furthermore, the protocol allowed additional analyses to be conducted to prioritise food-web research and monitoring efforts. For example, we applied the approach to assess how overall abundance balance between trophic levels might change in scenarios of loss of stability. Note that the approach could be easily adapted to make the analysis based on trophic guilds specifically defined for the study area (European Commission No. 19, 2022). We found that for the food-web modelled, impacts on dominant species, but not fisheries, could trigger a series of cascading effects leading to meaningful changes in food-web shape. Specifically, the model predicted a transition from a pyramidal trophic structure to a rhomboidal form, a pattern that approximates features reported for eutrophic systems (e.g. continental shelves) rather than oligotrophic open-ocean food-webs (Link et al., 2015). In the absence of empirical studies validating model outcomes, these results can serve to highlight research priorities. In this case, it seems relevant to better understand i) how the effects of predation-

release propagate through the food-web, ii) the role of zooplankton in the trophic structure and the consequences associated to their collapse and iii) the factors and mechanisms that most likely drive stability loss in the modelled system. Moreover, we applied the approach to identify the biological components that are most likely affected in scenarios of stability loss. The results, which are supported by the best available knowledge on the effects of perturbation of food-web stability, may be useful to policy makers in designing monitoring programmes.

5. Conclusion

This study reconciles ecological theory with statistical modelling as a methodological standard to assist the assessment and monitoring of the state of food-webs for achieving healthy and stable ecological states. In doing so, it adds to the ongoing scientific efforts in developing novel quantitative approaches for identifying indicators and thresholds suitable at detecting significant state changes and informing ecosystem-based management (e.g. Dakos et al., 2012, Fu et al., 2019, Libralato et al., 2019, Otto et al., 2018, Queirós et al., 2016, Rossberg et al., 2017, Samhouri et al., 2009, Samhouri et al., 2017). The key question to be further addressed is therefore: what steps are needed to implement the food-web stability information derived from the models? We recognise the importance of evaluating the proposed methodology in additional trophic models to determine its general validity, both in terms of indicators and the threshold identified as best for assessing food-web stability. Furthermore, it would be crucial to carry out real manipulations in natural systems or construct mesocosms (Stewart et al., 2013) to experimentally validate the results obtained from the modelling approach. By combining modelling and experimental approaches, a more comprehensive understanding of marine food-web responses to natural processes and human activities, as well as impacts on food-web stability, can be achieved (Chown, 2020, Nagelkerken et al., 2020) to support monitoring of Descriptor 4 towards GES.

CRediT authorship contribution statement

Joana Brito: Conceptualization, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. **Simone Libralato:** Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing. **Telmo Morato:** Conceptualization, Methodology, Supervision, Validation, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecolind.2023.111411.

References

Ahrens, R.N., Walters, C.J., Christensen, V., 2012. Foraging arena theory. Fish and Fisheries 13 (1), 41–59.

Andersen, K.H., Pedersen, M., 2010. Damped trophic cascades driven by fishing in model marine ecosystems. Proc. R. Soc. B Biol. Sci. 277 (1682), 795–802.

Andersen, K.H., Beyer, J.E., Lundberg, P., 2009. Trophic and individual efficiencies of size-structured communities. Proc. R. Soc. B Biol. Sci. 276 (1654), 109–114.

Anh, P.V., Everaert, G., Goethals, P., Vinh, C.T., De Laender, F., 2015. Production and food-web efficiency decrease as fishing activity increases in a coastal ecosystem. Estuar. Coast. Shelf Sci. 165, 226–236.

Arroyo, N.-L., Georges S., Pauline V., Lucía L.-L., Nathalie N., Francois L.L., Tarek H., Izaskun P. Towards coherent GES assessments at sub-regional level: signs of fisheries expansion processes in the Bay of Biscay using an OSPAR food-web indicator, the mean trophic level. ICES Journal of Marine Science 76, no. 6 (2019): 1543-1553.

Bartley, T.J., McCann, K.S., Bieg, C., Cazelles, K., Granados, M., Guzzo, M.M., MacDougall, A.S., Tunney, T.D., McMeans, B.C., 2019. Food-web rewiring in a changing world. Nat. Ecol. Evol. 3 (3), 345–354.

Bentley, J.W., Hines, D., Borrett, S., Serpetti, N., Fox, C., Reid, D.G., Heymans, J.J., 2019. Diet uncertainty analysis strengthens model-derived indicators of food-web structure and function. Ecol. Ind. 98, 239–250.

Bentley, J.W., Serpetti, N., Fox, C.J., Heymans, J.J., Reid, D.G., 2020. Retrospective analysis of the influence of environmental drivers on commercial stocks and fishing opportunities in the Irish Sea. Fish. Oceanogr. 29 (5), 415–435.

Branch, T.A., Watson, R., Fulton, E.A., Jennings, S., McGilliard, C.R., Pablico, G.T., Ricard, D., Tracey, S.R., 2010. The trophic fingerprint of marine fisheries. Nature 468 (7322), 431–435.

Carpenter, S.R., Ludwig, D., Brock, W.A., 1999. Management of eutrophication for lakes subject to potentially irreversible change. Ecol. Appl. 9 (3), 751–771.

Casini, M., Hjelm, J., Molinero, J.C., Lövgren, J., Cardinale, M., Bartolino, V., Belgrano, A., Kornilovs, G., 2009. Trophic cascades promote threshold-like shifts in pelagic marine ecosystems. Proc. Natl. Acad. Sci. 106 (1), 197–202.

Celić, I., Libralato, S., Scarcella, G., Raicevich, S., Marćeta, B., Solidoro, C., 2018. Ecological and economic effects of the landing obligation evaluated using a quantitative ecosystem approach: a Mediterranean case study. ICES J. Mar. Sci. 75 (6), 1992–2003.

Chagaris, D., Drew, K., Schueller, A., Cieri, M., Brito, J., Buchheister, A., 2020. Ecological reference points for Atlantic menhaden established using an ecosystem model of intermediate complexity. Front. Mar. Sci. 7, 606417.

Chown, S.L., 2020. Marine food-webs destabilized. Science 369 (6505), 770–771.

Christensen, V., 1995. Ecosystem maturity—towards quantification. Ecol. Model. 77 (1), 3–32.

Christensen, V., Walters, C.J., 2004. Ecopath with Ecosim: methods, capabilities and limitations. Ecol. Modell. 172 (2–4), 109–139.

Christensen, V., Walters, C.J., Pauly, D., Forrest, R., 2008. Ecopath with Ecosim version 6 user guide. Lenfest Ocean Futures Project 235.

Costanza, R., Mageau, M., 1999. What is a healthy ecosystem? Aquat. Ecol. 33, 105–115.

Dakos, V., Carpenter, S.R., Brock, W.A., Ellison, A.M., Guttal, V., Ives, A.R., Kéfi, S., Livina, V., Seekell, D.A., van Nes, E.H., Scheffer, M., 2012. Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. PLoS One 7 (7), e41010.

Donohue, I., Petchey, O.L., Montoya, J.M., Jackson, A.L., McNally, L., Viana, M., Healy, K., Lurgi, M., O'Connor, N.E., Emmerson, M.C., 2013. On the dimensionality of ecological stability. Ecol. Lett. 16 (4), 421–429.

Donohue, I., Hillebrand, H., Montoya, J.M., Petchey, O.L., Pimm, S.L., Fowler, M.S., Healy, K., Jackson, A.L., Lurgi, M., McClean, D., O'Connor, N.E., 2016. Navigating the complexity of ecological stability. Ecol. Lett. 19 (9), 1172–1185.

Du Pontavice, H., Gascuel, D., Reygondeau, G., Maureaud, A., Cheung, W.W., 2020. Climate change undermines the global functioning of marine food-webs. Glob. Chang. Biol. 26 (3), 1306–1318.

Dubois, M., Gascuel, D., Coll, M., Claudet, J., 2019. Recovery debts can be revealed by ecosystem network-based approaches. Ecosystems 22, 658–676.

Durante, L., Wing, S., Ingram, T., Sabadel, A., Shima, J., 2022. Changes in trophic structure of an exploited fish community at the centennial scale are linked to fisheries and climate forces. Sci. Rep. 12 (1), 4309.

Eddy, T.D., Bernhardt, J.R., Blanchard, J.L., Cheung, W.W., Colléter, M., Du Pontavice, H., Fulton, E.A., Gascuel, D., Kearney, K.A., Petrik, C.M., Roy, T., 2021. Energy flow through marine ecosystems: confronting transfer efficiency. Trends Ecol. Evol. 36 (1), 76–86.

Elith*, J., H. Graham*, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann, F., R. Leathwick, J., Lehmann, A. and Li, J., 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography, 29(2), pp.129-151.

European Commission No. 19, 2022. Article 8 MSFD, May 2022.

European Union, Commission Decision (EU) 2017/848 of 17 May 2017 laying down criteria and methodological standards on good environmental status of marine waters and specifications and standardized methods for monitoring and assessment, and repealing Decision 2010/477/EU. L 125/43, 2017.

FAO, 2009. Report of the Technical Consultation on International Guidelines for the Management of Deep-sea Fisheries in the High Seas, Rome. 4–8 February and 25–29 August 2008. FAO Fisheries and Aquaculture Report, pp. 881–886.

Fath, B.D., Asmus, H., Asmus, R., Baird, D., Borrett, S.R., de Jonge, V.N., Ludovisi, A., Niquil, N., Scharler, U.M., Schückel, U., Wolff, M., 2019. Ecological network analysis metrics: the need for an entire ecosystem approach in management and policy. Ocean Coast. Manag. 174, 1–14.

Fu, C., Xu, Y., Bundy, A., Grüss, A., Coll, M., Heymans, J.J., Fulton, E.A., Shannon, L., Halouani, G., Velez, L., Akoglu, E., 2019. Making ecological indicators management ready: assessing the specificity, sensitivity, and threshold response of ecological indicators. Ecol. Ind. 105, 16–28.

Geary, W.L., Bode, M., Doherty, T.S., Fulton, E.A., Nimmo, D.G., Tulloch, A.I., Tulloch, V.J., Ritchie, E.G., 2020. A guide to ecosystem models and their environmental applications. Nat. Ecol. Evol. 4 (11), 1459–1471.

Grimm, V., Wissel, C., 1997. Babel, or the ecological stability discussions: an inventory and analysis of terminology and a guide for avoiding confusion. Oecologia 109, 323–334.

Hale, S.S., Heltshe, J.F., 2008. Signals from the benthos: development and evaluation of a benthic index for the nearshore Gulf of Maine. Ecol. Ind. 8 (4), 338–350.

Heymans, J.J., Coll, M., Libralato, S., Morissette, L., Christensen, V., 2014. Global patterns in ecological indicators of marine food-webs: a modelling approach. PLoS One 9 (4), e95845.

Heymans, J.J., Coll, M., Link, J.S., Mackinson, S., Steenbeek, J., Walters, C., Christensen, V., 2016. Best practice in Ecopath with Ecosim food-web models for ecosystem-based management. Ecol. modell. 331, 173–184.

Heymans, J.J., 2003. Ecosystem models of Newfoundland and Southeastern Labrador: additional information and analyses for" back to the future".

Holling, C.S., 1973. Resilience and stability of ecological systems. Annu. Rev. Ecol. Syst. 4 (1), 1–23.

Ives, A.R., Carpenter, S.R., 2007. Stability and Diversity of Ecosystems. Science 317 (5834), 58–62.

Kauhala, K., Korpinen, S., Lehtiniemi, M., Raitaniemi, J., 2019. Reproductive rate of a top predator, the grey seal, as an indicator of the changes in the Baltic food-web. Ecol. Ind. 102, 693–703.

Korpinen, S., Uusitalo, L., Nordström, M.C., Dierking, J., Tomczak, M.T., Haldin, J., Opitz, S., Bonsdorff, E., Neuenfeldt, S., 2022. Food-web assessments in the Baltic Sea: models bridging the gap between indicators and policy needs. Ambio 51 (7), 1687–1697.

Kortsch, S., Primicerio, R., Fossheim, M., Dolgov, A.V., Aschan, M., 2015. Climate change alters the structure of arctic marine food-webs due to poleward shifts of boreal generalists. Proc. R. Soc. B Biol. Sci. 282 (1814), 20151546.

Libralato, S., Solidoro, C., 2010. Comparing methods for building trophic spectra of ecological data. ICES J. Mar. Sci. 67 (3), 426–434.

Libralato, S., Christensen, V., Pauly, D., 2006. A method for identifying keystone species in food-web models. Ecol. Model. 195 (3–4), 153–171.

Libralato, S., Pranovi, F., Zucchetta, M., Monti, M.A., Link, J.S., 2019. Global thresholds in properties emerging from cumulative curves of marine ecosystems. Ecol. Ind. 103, 554–562.

Link, J.S., Pranovi, F., Libralato, S., Coll, M., Christensen, V., Solidoro, C., Fulton, E.A., 2015. Emergent properties delineate marine ecosystem perturbation and recovery. Trends Ecol. Evol. 30 (11), 649–661.

Machado, I., Costa, J.L., Cabral, H., 2021. Response of food-webs indicators to human pressures, in the scope of the marine strategy framework directive. Front. Mar. Sci. 8, 699566.

Macmillan, N.A., Creelman, C.D., 2004. Detection theory: A user's guide. Psychology press.

Mageau, M.T., Costanza, R., Ulanowicz, R.E., 1998. Quantifying the trends expected in developing ecosystems. Ecol. Model. 112 (1), 1–22.

Maureaud, A., Gascuel, D., Colléter, M., Palomares, M.L., Du Pontavice, H., Pauly, D., Cheung, W.W., 2017. Global change in the trophic functioning of marine food-webs. PLoS One 12 (8), e0182826.

McCann, K.S., 2000. The diversity–stability debate. Nature 405 (6783), 228–233.

McCann, K.S., 2011. Food-webs (MPB-50). In Food-webs (MPB-50). Princeton University Press.

McNaughton, S.J., Wolf, L.L., 1970. Dominance and the Niche in Ecological Systems: dominance is an expression of ecological inequalities arising out of different exploitation strategies. Science 167 (3915), 131–139.

Möllmann, C., Folke, C., Edwards, M., Conversi, A., 2015. Marine regime shifts around the globe: theory, drivers and impacts. Philos. Trans. R. Soc., B 370 (1659), 20130260.

Morato, T., Lemey, E., Menezes, G., Pham, C.K., Brito, J., Soszynski, A., Pitcher, T.J., Heymans, J.J., 2016. Food-web and ecosystem structure of the open-ocean and deep-sea environments of the Azores. NE Atlantic. Frontiers in Marine Science 3, 245.

Moreno-Mateos, D., Barbier, E.B., Jones, P.C., Jones, H.P., Aronson, J., López-López, J.A., McCrackin, M.L., Meli, P., Montoya, D., Rey Benayas, J.M., 2017. Anthropogenic ecosystem disturbance and the recovery debt. Nat. Commun. 8 (1), 14163.

Murtaugh, P.A., 1996. The statistical evaluation of ecological indicators. Ecol. Appl. 6 (1), 132–139.

Nagelkerken, I., Goldenberg, S.U., Ferreira, C.M., Ullah, H., Connell, S.D., 2020. Trophic pyramids reorganize when food-web architecture fails to adjust to ocean change. Science 369 (6505), 829–832.

Novaglio, C., Smith, A.D., Frusher, S., Ferretti, F., 2020. Identifying historical baseline at the onset of exploitation to improve understanding of fishing impacts. Aquat. Conserv. Mar. Freshwat. Ecosyst. 30 (3), 475–485.

Nyström, M., Norström, A.V., Blenckner, T., de la Torre-Castro, M., Eklöf, J.S., Folke, C., Österblom, H., Steneck, R.S., Thyresson, M., Troell, M., 2012. Confronting feedbacks of degraded marine ecosystems. Ecosystems 15, 695–710.

Odum, E.P., 1985. Trends expected in stressed ecosystems. Bioscience 35 (7), 419–422.

Olliaro, P., Torreele, E., 2021. Managing the risks of making the wrong diagnosis: first, do no harm. Int. J. Infect. Dis. 106, 382–385.

Otto, S.A., Kadin, M., Casini, M., Torres, M.A., Blenckner, T., 2018. A quantitative framework for selecting and validating food-web indicators. Ecol. Ind. 84, 619–631.

Paine, R.T., 1969. A note on trophic complexity and community stability. Am. Nat. 103 (929), 91–93.

Pauly, D., 1995. Anecdotes and the shifting baseline syndrome of fisheries. Trends Ecol. Evol. 10 (10), 430.

Pimm, S.L., 1984. The complexity and stability of ecosystems. Nature 307 (5949), 321–326.

Piroddi, C., Teixeira, H., Lynam, C.P., Smith, C., Alvarez, M.C., Mazik, K., Andonegi, E., Churilova, T., Tedesco, L., Chifflet, M., Chust, G., 2015. Using ecological models to assess ecosystem status in support of the european marine strategy framework directive. Ecol. Ind. 58, 175–191.

Planque, B., 2016. Projecting the future state of marine ecosystems, "la grande illusion"? ICES J. Mar. Sci. 73 (2), 204–208.

Polovina, J.J., 1984. Model of a coral reef ecosystem: I. The ECOPATH model and its application to French Frigate Shoals. Coral Reefs 3, 1–11.

Queirós, A.M., Strong, J.A., Mazik, K., Carstensen, J., Bruun, J., Somerfield, P.J., Bruhn, A., Ciavatta, S., Flo, E., Bizsel, N., Özaydinli, M., 2016. An objective framework to test the quality of candidate indicators of good environmental status. Front. Mar. Sci. 3, 73.

Rice, J., 2003. Environmental health indicators. Ocean Coast. Manag. 46 (3–4), 235–259.

Rogers, S., Casini, M., Cury, P., Heath, M., Irigoien, X., Kuosa, H., Scheidat, M., Skov, H., Stergiou, K., Trenkel, V., Wikner, J., 2010. Marine Strategy Framework Directive-Task Group 4 Report Food-webs. European Commission Joint Research Centre, ICES.

Rombouts, I., Beaugrand, G., Artigas, L.F., Dauvin, J.C., Gevaert, F., Goberville, E., Kopp, D., Lefebvre, S., Luczak, C., Spilmont, N., Travers-Trolet, M., 2013a. Evaluating marine ecosystem health: case studies of indicators using direct observations and modelling methods. Ecol. Ind. 24, 353–365.

Rombouts, I., Beaugrand, G., Fizzala, X., Gaill, F., Greenstreet, S.P.R., Lamare, S., 2013b. Food-web indicators under the Marine Strategy Framework Directive: from complexity to simplicity? Ecol. Ind. 29, 246–254.

Rossberg, A.G., Uusitalo, L., Berg, T., Zaiko, A., Chenuil, A., Uyarra, M.C., Borja, A., Lynam, C.P., 2017. Quantitative criteria for choosing targets and indicators for sustainable use of ecosystems. Ecol. Ind. 72, 215–224.

Safi, G., Giebels, D., Arroyo, N.L., Heymans, J.J., Preciado, I., Raoux, A., Schückel, U., Tecchio, S., de Jonge, V.N., Niquil, N., 2019. Vitamine ENA: a framework for the development of ecosystem-based indicators for decision makers. Ocean Coast. Manag. 174, 116–130.

Saint-Béat, B., Baird, D., Asmus, H., Asmus, R., Bacher, C., Pacella, S.R., Johnson, G.A., David, V., Vézina, A.F., Niquil, N., 2015. Trophic networks: How do theories link ecosystem structure and functioning to stability properties? A review. Ecol. Ind. 52, 458–471.

Samhouri, J.F., Levin, P.S., Harvey, C.J., 2009. Quantitative evaluation of marine ecosystem indicator performance using food-web models. Ecosystems 12, 1283–1298.

Samhouri, J.F., Andrews, K.S., Fay, G., Harvey, C.J., Hazen, E.L., Hennessey, S.M., Holsman, K., Hunsicker, M.E., Large, S.I., Marshall, K.N., Stier, A.C., 2017. Defining ecosystem thresholds for human activities and environmental pressures in the California Current. Ecosphere 8 (6), e01860.

Scheffer, M., Carpenter, S., Foley, J.A., Folke, C., Walker, B., 2001. Catastrophic shifts in ecosystems. Nature 413 (6856), 591–596.

Scheffer, M., Carpenter, S.R., 2003. Catastrophic regime shifts in ecosystems: linking theory to observation. Trends in ecology evolution 18 (12), 648–656.

Shephard, S., Greenstreet, S.P., Piet, G.J., Rindorf, A., Dickey-Collas, M., 2015. Surveillance indicators and their use in implementation of the Marine Strategy Framework Directive. ICES J. Mar. Sci. 72 (8), 2269–2277.

Shin, Y.J., Bundy, A., Shannon, L.J., Simier, M., Coll, M., Fulton, E.A., Link, J.S., Jouffre, D., Ojaveer, H., Mackinson, S., Heymans, J.J., 2010. Can simple be useful

and reliable? Using ecological indicators to represent and compare the states of marine ecosystems. ICES J. Mar. Sci. 67 (4), 717–731.

Steenbeek, J., Corrales, X., Platts, M., Coll, M., 2018. Ecosampler: a new approach to assessing parameter uncertainty in Ecopath with Ecosim. SoftwareX 7, 198–204.

Stewart, R.I., Dossena, M., Bohan, D.A., Jeppesen, E., Kordas, R.L., Ledger, M.E., Meerhoff, M., Moss, B., Mulder, C., Shurin, J.B., Suttle, B., 2013. Mesocosm experiments as a tool for ecological climate-change research. Adv. Ecol. Res. 48, 71–181.

Storch, L.S., Glaser, S.M., Ye, H., Rosenberg, A.A., 2017. Stock assessment and end-to-end ecosystem models alter dynamics of fisheries data. PLoS One 12 (2), e0171644.

Tam, J.C., Link, J.S., Rossberg, A.G., Rogers, S.I., Levin, P.S., Rochet, M.J., Bundy, A., Belgrano, A., Libralato, S., Tomczak, M., Van De Wolfshaar, K., 2017. Towards ecosystem-based management: identifying operational food-web indicators for marine ecosystems. ICES J. Mar. Sci. 74 (7), 2040–2052.

Thiele, C., Hirschfeld, G., 2021. cutpointr: improved estimation and validation of optimal cutpoints in R. J. Stat. Softw. 98 (11), 1–27. https://doi.org/10.18637/jss.v098.i11.

Torres, M.A., Casini, M., Huss, M., Otto, S.A., Kadin, M., Gårdmark, A., 2017. Food-web indicators accounting for species interactions respond to multiple pressures. Ecol. Ind. 77, 67–79.

Tverin, M., Esparza-Salas, R., Strömberg, A., Tang, P., Kokkonen, I., Herrero, A., Kauhala, K., Karlsson, O., Tiilikainen, R., Vetemaa, M., Sinisalo, T., 2019. Complementary methods assessing short and long-term prey of a marine top predator-Application to the grey seal-fishery conflict in the Baltic Sea. PLoS One 14 (1), e0208694.

Ulanowicz, R.E., 2004. Quantitative methods for ecological network analysis. Comput. Biol. Chem. 28 (5–6), 321–339.

Ullah, H., Nagelkerken, I., Goldenberg, S.U., Fordham, D.A., 2018. Climate change could drive marine food-web collapse through altered trophic flows and cyanobacterial proliferation. PLoS Biol. 16 (1), e2003446.

Van Meerbeek, K., Jucker, T., Svenning, J.C., 2021. Unifying the concepts of stability and resilience in ecology. J. Ecol. 109 (9), 3114–3132.

Van Nes, E.H., Scheffer, M., 2007. Slow recovery from perturbations as a generic indicator of a nearby catastrophic shift. The Am. Nat. 169 (6), 738–747.

Verling, E., Bartilotti, C., Hollatz, C., Tuaty-Guerra, M., Lobo-Arteaga, J., O'Higgins, T., 2023. Applying risk-based approaches to implementation of the Marine Strategy Framework Directive in the North-East Atlantic: Learning lessons and moving forward. Mar. Policy 153, 105667.

Walters, C., Christensen, V., Pauly, D., 1997. Structuring dynamic models of exploited ecosystems from trophic mass-balance assessments. Rev. Fish Biol. Fish. 7 (2), 139.