# Sustainable fishing can lead to improvements in marine ecosystem status: an ensemble-model forecast of the North Sea ecosystem 

Spence, Michael; Griffiths, Christopher ; Waggitt, James; Bannister, Hayley; Thorpe, Robert; Rossberg, Axel; Lynam, Christopher

# Marine Ecology Progress Series 

DOI:
https://doi.org/10.3354/meps13870
Published: 09/12/2021

Peer reviewed version

Cyswllt i'r cyhoeddiad / Link to publication

Dyfyniad o'r fersiwn a gyhoeddwyd / Citation for published version (APA):
Spence, M., Griffiths, C., Waggitt, J., Bannister, H., Thorpe, R., Rossberg, A., \& Lynam, C. (2021). Sustainable fishing can lead to improvements in marine ecosystem status: an ensemblemodel forecast of the North Sea ecosystem. Marine Ecology Progress Series, 680, 207-221. https://doi.org/10.3354/meps13870

## Hawliau Cyffredinol / General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal?

[^0][^1]
#### Abstract

To effectively implement Ecosystem Based Fisheries Management, tools are needed that are capable of exploring the likely consequences of potential management action on the whole ecosystem. Quantitative modelling tools can be used to explore how ecosystems might respond to potential management measures, but no one model can reliably forecast all aspects of future change. To build a robust basis for management advice, a suite of models can be used, but the interpretation of the joint output of multiple models can be difficult. We employ a newly developed ensemble approach to integrate five different ecosystem models and estimate changes in ecosystem state within a single probabilistic forecast. We provide evidence on the response of ecosystem state (measured using ecological indicators relating to plankton, fish and top predators) to potential fisheries management scenarios. We demonstrate that if future fishing mortality is consistent with MSY policy, the North Sea fish community will recover in terms of its size-structure and species composition. However, there is currently large uncertainty in trends of future fish biomass, plankton and top predators. We conclude that a) this ensemble approach can be applied directly to policy-relevant questions and add value for decision-makers as multiple aspects of uncertainty are considered; b) future research should prioritise improvements in model skill via a reduction in uncertainty surrounding biomass estimates; c) fisheries management that leads to sustainable fishing levels can be considered appropriate for two crucial aspects of fish biodiversity; species composition and size structure.


Key words: ecosystem models, ensemble modelling, indicators, maximum sustainable yield, multispecies models, North Sea, uncertainty.

## 1. Introduction

Many national governments have a policy to implement an ecosystem-based approach to fisheries management (EBFM); and the parties to the Convention on Biological Diversity (www.cbd.int) have agreed Strategic Goals with specific targets to "Reduce the direct pressures on biodiversity and promote sustainable use" and "Improve the status of biodiversity by safeguarding ecosystems, species and genetic diversity". An overarching aim is to reduce direct pressures on biodiversity and promote sustainable use of our seas. To measure progress towards these goals, ecological indicators have been developed for national (e.g., HM Government 2019) and regional (OSPAR 2017a b c d, HELCOM 2018) assessments including indicators of pressure (e.g., fishing mortality, noise, contaminants) and state (species, habitats and communities; McQuatters-Gollop et al. 2019). In the European Union, indicator-based assessments also support the policy objectives of the Marine Strategy Framework Directive (European Commission 2008) and the reformed Common Fisheries Policy (European Union 2015). In the UK, indicator assessments support the UK Marine Strategy Regulations (HM Government 2010) and the ecosystem-based approach is embedded within the UK Fisheries Act (HM Government 2020), with an explicit objective of achieving good environmental status (GES), where the seas are clean, healthy, safe, productive and biologically diverse (2008/56/EC). To implement EBFM, the indicator assessments are followed by a review of monitoring required to capture human impacts on the marine environment and recommendations for a programme of measures to achieve GES. This process is called the indicator assessment cycle (Lynam et al. 2016).

Fishing is a major pressure on marine ecosystems (Jennings \& Kaiser 1998, Pauly et al. 1998) and numerous indicators have been developed to describe how changes in state have, and
will, respond to fisheries management scenarios (Fulton et al. 2005, Collie et al. 2013, Coll et al. 2016, Rossberg et al. 2017, Fu et al. 2019). Some indicators consider differences in habitat, e.g., relative benthic status (Hiddink et al. 2020), whereas others consider the effect that fishing might have on charismatic species of conservation interest, e.g., seabird breeding success (Eerkes-Medrano et al. 2017). Here, we are particularly interested in changes in ecosystem structure and function, specifically functional group indicators, as they provide integrative measures that can capture the wider effects of fisheries on ecosystems (including indirect impacts that may propagate up through the food web, e.g., to piscivorous seabirds and marine mammals) and of ecosystem change on fisheries (e.g., through changes in the biomass of plankton at the base of food web; Tam et al. 2017). Community level indicators of size structure (Bianchi et al. 2000, Greenstreet et al. 2012, Brose et al. 2017, Tu et al. 2018) and species composition (Fisher et al. 2010, Houle et al. 2012, Bell et al. 2018) have previously been shown to respond to fishing pressure. These ecological indicators have been incorporated into assessments by OSPAR and, within the UK, they are used to monitor changes in food webs and biodiversity, alongside stock-specific information for commercial fish and shellfish (see UK Marine Online Assessment Tool, UKMMAS 2019). We consider a selection of these indicators of GES for which current published ecosystem models are able to inform on: i.e., the biomass of zooplankton, the ratio of phytoplankton to zooplankton biomass, the biomass of fish and top predators (birds and mammals), along with measures of species composition and size structure within the fish community.

Within the indicator assessment cycle, modelling should be employed to assess the performance of indicators in relation to pressure and forecast any likely change in state following the implementation of existing and/or planned management measures (Lynam et al. 2016). For example, limiting fishing opportunities (catch quotas or days at sea) in order to
achieve maximum sustainable yield (MSY), the highest yield achievable in the long-term while maintaining reproductive capacity of the stock (ICES 2021). Several possible ecosystem models could, and have, been used for this task. For example, Blanchard et al. (2014) used a size-based model to investigate how length-based indicators (including the Large Fish Indicator, LFI) respond to past and future fishing pressure. Speirs et al. (2016) demonstrated that fishing at MSY could facilitate recovery in LFI to levels that represent GES in the North Sea, where the indicator target was defined by Greenstreet et al. (2011). However, Queirós et al. (2018) also found through spatial and size-based modelling that any recovery in LFI would be limited by climate effects on fish body size. Additionally, Ecopath with Ecosim has been employed to test the relevance of trophic indicators (including the slope of the biomass spectrum) to EBFM in numerous ecosystems (Bourdaud et al. 2016). Despite such efforts, it is important to note that individual ecosystem models, despite being built with similar processes and fitted to similar data, can often produce conflicting predictions (Plagányi 2007, Collie et al. 2016, Tittensor et al. 2017, Rosenberg et al. 2018). Consequently, the predicted outcome of any management decision under consideration is sensitive to the model chosen. This sensitivity to model choice, coupled with the inherent uncertainty in modelling the complex dynamics of whole ecosystems has, in part, led to a lack of uptake of ecosystem models in the advice process (Hyder et al. 2015).

When giving advice it is important to quantify uncertainty in predictions in a robust manner (Harwood \& Stokes 2003). Uncertainty within individual ecosystem models has previously been explored (e.g., Thorpe et al. 2015, Spence et al. 2016, 2021, Mackinson et al. 2018), however, these attempts are likely to underestimate the uncertainty, as uncertainty across models is often ignored. Management advice should ideally consider the outputs of a suite of models, however, combining them is not as easy as averaging over their outputs or ascribing
a simple weight to each as we do not expect the discrepancies of each of the models to be independent (Rougier 2007, Christiansen 2020). Further, each model's relative utility is often question-sensitive (Dickey-Collas et al. 2014), and no such model will be uniformly better than the rest (Chandler 2013). To counteract this problem, Spence et al. (2018) developed an ensemble model that synthesises empirical and modelling studies to give a single probabilistic forecast with robust quantification of uncertainty, however, this approach is yet to be employed to investigate indicators of pressure or ecological state. To supplement our approach, we also assess 'model skill' (here using information), defined as the inverse of uncertainty, as a metric. This is an emergent feature in the recent literature (e.g., Olsen et al. 2016, Geary et al. 2020) and provides a useful narrative via which this comparison of uncertainty can be made.

In this study, we address an explicit policy question that follows from the OSPAR 2017 (OSPAR 2017a b c d, HELCOM 2018) and UK 2019 assessments of biodiversity (HM Government 2019), namely: will fisheries management measures that aim to maintain productive commercial fisheries also lead to improvements across the ecosystem? The long-term impact of fishing on an ecosystem includes the depletion in the proportion of large or old individuals (Christensen 1996, Pauly et al. 1998, Greenstreet et al. 2011), a reduction in the biomass of fished populations (Palomares et al. 2020) and subsequent declines in the abundance of apex predators (mammals, birds, large fish and elasmobranch species) that are dependent on the availability of prey (Lynam et al. 2017). A recovery in size-structure (using typical length and/or LFI) is considered important for food web functioning as large individuals are typically effective spawners (Barneche et al. 2018) and top predators play an important role in stabilising ecosystems (Casini et al. 2012, Lynam et al. 2017, Landi et al. 2018). A simultaneous recovery in size-structure and species composition, as measured by mean maximum length
(MML), is considered a strong measure of improved biodiversity since species with a high maximum body length are typically slow-growing and most vulnerable to fishing impacts (Andersen 2019). So, a question that arises from managers seeking an improved ecosystem state is: will fish communities (as measured by their size-structure and species composition) and the biomass of birds and mammals improve as stocks are fished at their target mortalities, or will more restrictive mortality targets be required? Similarly, a related question posed is: will an increase in fish biomass lead to a depletion in zooplankton production or is the ecosystem productive enough to allow species to recover? To provide an integrative understanding of the likely impacts of fisheries management on the North Sea ecosystem, we investigate changes in the aforementioned indicators of GES through the ensemble modelling approach of Spence et al. (2018). We consider a range of management scenarios (including fishing at single species or multi-species targets for MSY), five different ecosystem models (Ecopath with EcoSim, mizer, FishSUMS, StrathE2E and LeMans) and an empirical analysis of the indicators of interest, all within a single ensemble model. This approach allows us to (a) make use of information from numerous ecosystems models; (b) identify the likely outcomes of each fisheries management scenario with robust and quantifiable uncertainty; (c) discuss how changes in fishing pressure will influence ecosystem structure and function; (d) assess the model skill of the suite of ecosystem models and make targeted recommendations for model development.

## 2. Materials \& Methods

### 2.1. North Sea

The North Sea is a semi-enclosed sea situated on the continental shelf of North Western Europe. Many offshore fisheries operate in the North Sea and are regulated in accordance
with the EU Common Fisheries Policy (European Union 2015) and by Norway and the UK through coastal state agreements. The environmental policy for the North Sea is managed by national governments and OSPAR (e.g., OSPAR 2009). Demersal fishing effort (beam trawling and demersal otter trawling, Couce et al. 2020) has been reduced since the 2002 Common Fisheries Policy reforms (European Union 2015). Since then, a recovering trend in the demersal fish community has been observed (Engelhard et al. 2015), however, there is substantially less food (through a reduction in discards) for surface feeding and scavenging seabirds (Sherley et al. 2020).

### 2.2. Indicators of state

The following indicators of GES were considered:

1. Total biomass of marine birds and mammals ( $B M$, tonnes). $B M$ will respond to fishing indirectly via shifts in food availability and/or directly due to mortality caused by human activities.
2. Total biomass of the demersal fish community (TFB, tonnes). TFB will be impacted directly by ecosystem productivity and fishing pressure.
3. Size-structure of the demersal fish community was measured using the large fish indicator (LFI), which is the proportion of biomass in the community that is composed of individuals larger than a set threshold: here 40 cm , i.e.,

$$
\begin{equation*}
L F I=\frac{\sum_{i=1}^{n} B_{i} I\left(l_{i} \geq 40\right)}{\sum_{i=1}^{n} B_{i}} \tag{1}
\end{equation*}
$$

This indicator has been used for many years in the ICES and OSPAR communities (ICES 2009, OSPAR 2009, 2017c) and was developed to be responsive to fishing pressure. LFI will decline as large individuals are fished out of the ecosystem (Greenstreet et al. 2011).
4. Size-structure within the whole fish community was investigated through the Typical Length indicator (TyL, cm, the geometric mean length of fish weighted by biomass; OSPAR 2017d), which has been suggested to have more favourable statistical qualities than the ratio indicator LFI (ICES 2014). The typical length indicator is

$$
\begin{equation*}
T y L=\exp \left[\frac{\sum_{i=1}^{n} B_{i} \log \left(l_{i}\right)}{\sum_{i=1}^{n} B_{i}}\right] \tag{2}
\end{equation*}
$$

where $B_{i}$ is the biomass and $l_{i}$ is the length ( cm ) of the $i$ th fish in a sample of $n$ fish. Ty $L$ will rise and fall as the size-structure responds to fishing pressure and other pressures including climate change impacts (Houle et al. 2012, Tu et al. 2018).
5. Species-composition within the whole fish community was measured using the Mean Maximum Length indicator (MML, cm). The Mean Maximum Length indicator is calculated as:

$$
\begin{equation*}
M M L=\frac{\sum_{j=1}^{m} B_{j} L_{i n f, j}}{\sum_{j=1}^{m} B_{j}} \tag{3}
\end{equation*}
$$

where $B_{j}$ is the biomass of each species $j$ for a community of $m$ species and $L_{i n f, j}$ is the growth-model asymptote known as length at infinity. MML will decrease as communities are fished down (Houle et al. 2012). The rationale for this is that large, slow-growing, slow to mature species are more sensitive to fishing pressure impacts than smaller-bodied, rapidly growing, fast to mature species. Under the "fishing-down-the-food-web" hypothesis (Christensen 1996, Pauly et al. 1998), larger bodied species (typically more valuable species) are targeted preferentially by fisheries, which would lead to a direct decline in $M M L$ and the size-structure of the fish community. For the model studies $L_{i n f, j}$ is specific to each model and for the empirical analysis estimates were taken from FishBase (Froese \& Pauly 2021).
6. The ratio of zooplankton to phytoplankton biomass ( $Z: P$, dimensionless). $Z: P$ will respond to top-down cascades when they are strong (Rossberg et al. 2019) and may also respond to climate change (OSPAR 2017a). We expect Z:P to respond indirectly to fishing pressure via the recruitment of fish larvae and the consumption of zooplankton by zooplanktivorous fish (Gorokhova et al. 2016).
7. Total zooplankton biomass ( $Z B$, tonnes). $Z B$ represents the transmission of bottom-up effects, e.g., related to nutrient concentrations, to zooplanktivorous fish (Gorokhova et al. 2016). We expect $Z B$ to respond directly to ecosystem productivity and indirectly to fishing pressure as the numbers of potential predators in the ecosystem change.

The historical change in each indicator was examined from 1986 to 2013 based on available data and/or empirical studies (see Section 2.3). From 2014 to 2050, each indicator was forecasted using all available ecosystem models (not all ecosystem models were capable of forecasting each indicator) and the fisheries management scenario under consideration.

### 2.3. Empirical studies

Continuous Plankton Recorder (CPR) data for zooplankton (abundance of copepods; individuals per n. $\mathrm{mi}^{2}$ ) and phytoplankton (total counts of diatoms and dinoflagellates; individuals per $\mathrm{n} . \mathrm{mi}^{2}$ ) were taken from the study of Lynam et al. (2013). Z:P was calculated by division of annual records for each component. $Z B$ was determined from abundance data for 29 taxa using wet weight factors and corrections for under-sampling of the CPR relative to vertical towed nets (WP-2), using the conversion factors of Pitois \& Fox (2006) (see Supplement 1, Table S1). We modelled the seasonality and long-term change in ZB using a generalized additive model (see Supplement 1).

The $M M L$ for all species, $T y L$ for all species, $T F B$ and $L F /$ were calculated from the International Bottom Trawl Survey (IBTS) data from Quarter 1 (January - March). Processed data were obtained from the quality assured OSPAR Groundfish Survey Monitoring and Assessment Data Product for the northeast Atlantic area (Greenstreet \& Moriarty 2017, Moriarty et al. 2017), which is based on national datasets uploaded to the ICES DATRAS database. Walker et al. (2017) developed catchability correction factors for species and size classes of fish in the IBTS and we apply them here to better estimate fish biomass and community indicators. A second order polynomial regression was fitted to log transformed MML, TyL and TFB to smooth the data and calculate uncertainty values (see Supplement 2 for more details). Similarly, a second order polynomial regression was fitted to the logistic transform of the LFI. $B M$ was taken from species distribution model predictions based on $\sim 2$ million kilometres of at-sea cetacean and seabird surveys between 1985 and 2015 (Waggitt et al. 2020; see Supplement 1).

### 2.4. Future management scenarios

The likely change in the ecosystem state (as measured by the indicators described in Section 2.2) was forecasted to the year 2050 under four different fishing management scenarios:

1. Status quo: continue to fish at the 2013 fishing mortality rates for all stocks (Table 1) as defined by Thorpe et al. (2017).
2. Maximum sustainable yield (MSY): fish 21 species at the fishing mortality rates recommend by ICES for obtaining MSY, while all other species are fished at 2013 levels. MSY is the maximum yield of a single species that can be sustainably caught and is calculated on a species by species basic. All values are taken from Thorpe et al. (2017).

See Supplement 3 for the specific definition of MSY and Table 1 for species-specific fishing mortality rates.
3. Nash equilibrium (Nash): fish 21 species at mortality rates according to the Nash equilibrium solution computed by Thorpe et al. (2017), while all other species are fished at 2013 levels. The Nash equilibrium is a potential multispecies MSY, where each of the individual species are simultaneously at their respective single species MSYs. See Supplement 3 for the specific definition of the Nash equilibrium and Table 1 for speciesspecific fishing mortality.
4. Closure (termed no fishing throughout): no fishing at all Each scenario was implemented from the year 2014. For more information on the definitions of each scenario please see Supplement 3.

### 2.5. Ecosystem models

Management effects on the indicators were predicted using the five ecosystem models that were available to us. They were:

1. Ecopath with EcoSim (EwE) is a whole ecosystem model that was first developed in 1984 by Polovina (1984) and has been updated since to include temporal (Ecosim) and spatial (Ecospace) dynamics (Christensen \& Walters 2004). Ecopath is currently used extensively across the globe to simulate historic changes in ecosystems and their response to pressures (Heymans et al. 2016). The EwE model used here is the North Sea model reviewed by ICES (2016). It contains > 10 fishing fleets and $>60$ functional groups, some of which are split into multiple age stanzas. We used parameter uncertainty from Mackinson et al. (2018). EwE was able to predict ZB, Z:P, BM and TFB from 1991 until 2050.
2. The multispecies size spectrum model (mizer) was developed to represent the size and abundance of all organisms from zooplankton to large fish predators in a size-structured food web. A proportion of the organisms are represented by species-specific traits and body size while others are represented solely by body size. The core of the model involves ontogenetic feeding and growth, mortality, and reproduction driven by size-dependent predation and maturation processes (Hartvig et al. 2011, Scott et al. 2014). Blanchard et al. (2014) developed and applied a version of mizer for the North Sea that is used here. Mizer was able to predict MML, TyL, LFI and TFB from 1986 until 2050.
3. FishSUMs (Speirs et al. 2010, 2016) represents the population dynamics of a set of key tropically-linked predator and prey species. For each species the state variables are biomass by length class. In discrete time steps the state variables are updated through increasing length, density-dependent mortality, and losses due to fishing and predation by explicitly modelled species, and seasonal reproduction. Additional food resources, not modelled at the species level, are characterised by three biomass spectra representing zooplankton, benthos, and "other fish". The model was initially configured for the North Sea with a set of nine structured species (Speirs et al. 2010), and subsequently extended to include plaice and saithe so as to include the eight most abundant demersal species that make up > 90\% of the North Sea biomass (Speirs et al. 2016). FishSUMs was able to predict MML, TyL, LFI and TFB from 1986 until 2050.
4. The Strathclyde end-to-end (StrathE2E) marine food web model is a whole ecosystem model designed to simulate regional scale, macroscopic top-down and bottom-up cascading trophic effects (Heath et al. 2014). The mathematical formulation is based on a network of coupled ordinary differential equations representing the entire food web in the water column and seabed sediments from nutrients and microbes though
zooplankton and fish, to birds and mammals, including the effects of advection, mixing and active vertical migrations. Living components are represented at low taxonomic resolution, focussing on fluxes of nitrogen between coarse functional groups, and simulating the general "shape" of the food web rather than the detail. StrathE2E was able to predict $Z B, Z: P$ and $B M$ from 1991 until 2050.
5. LeMans is a modified form of the length-based multispecies model initially developed by Hall et al. (2006) to represent the Georges Bank fish community, which was subsequently adapted for use in the North Sea by Rochet et al. (2011). The model represents 21 fish species in 32 equal length classes of around 5 cm each, spanning the full-size range of species represented into the model (Thorpe et al. 2015, Spence et al. 2020). LeMans was able to predict MML, TYL, LFI and TFB from 1986 until 2050.

Each model was run with historical fishing mortalities applied annually until 2013 and then the four future scenarios were implemented until 2050. From 2014-2050, fishing mortality rates were fixed according to the management strategies described above (see Section 2.3 and Table 1). A summary of each ecosystem model and their outputs with respect to this study can be found in Supplement 4.

### 2.6. Ensemble model

Following the approach of Spence et al. (2018), the five ecosystem models (see Section 2.5) were combined with empirical studies (see Section 2.3 ) in an ensemble model. The ensemble model is a statistical model that describes the structure of the suite of models and the empirical studies. It separates model discrepancies (or biases) between "shared discrepancies", i.e., discrepancies shared between models, and "individual discrepancies", which are specific to each model. It combines the model outputs and discrepancies with
probabilistic estimates from the empirical studies to generate a prediction of the true value of the indicators, with quantifiable uncertainty. A more thorough description of the ensemble model is provided in Supplement 5.

The ensemble model was fitted using a Bayesian framework in order to quantify the uncertainty (Spence et al. 2018). For prior specification see Supplement 5. The likelihood of the ensemble model is intractable and so we sampled from the posterior distribution using Markov Chain Monte Carlo (MCMC). Due to the high dimensionality and correlation of the uncertain parameter space, we fitted the model using No U-turn Hamiltonian Monte Carlo (Hoffman \& Gelman 2014) in the package Stan (version 2.21.2, Carpenter et al. 2017) in R (version 4.0.2, R Core Team 2020). The ensemble model was fitted by running the MCMC algorithm for 2000 iterations. The first 1000 iterations were discarded as a burn-in.

## 3. Results

### 3.1. Ecosystem model runs

The ability of each ecosystem model to follow the empirical studies and then forecast under an MSY scenario is illustrated in Figure 1 (see Supplement 6 for alternative scenarios). Only two of the models (EwE and StrathE2E) were able to estimate $Z B, Z: P$ and $B M$ and both were unable to replicate the historical trends. There is also substantial disagreement between models when forecasting, both in terms of trends and absolute values. StratheE2E predicted an increase in $Z B$ until 2050 whereas EwE did not. Similarly, $B M$ and $Z: P$ were estimated to increase (albeit slightly) by EwE, whereas StrathE2E estimated no change.

Three models (mizer, FishSUMS and LeMans) were able to estimate TyL and LFI. All three models followed historical trends, with both TyL and LFI increasing over 1986-2013. This increase was then accelerated, followed by a plateau in all three model forecasts. Alongside

TyL and LFI, mizer, fishSUMS and LeMans, as well as EwE, were able to estimate changes in species composition through $M M L$. Trends in $M M L$ during the historic period did vary between models, however, they all suggested some degree of increase. This increase of MML is analogous to the empirical studies despite large uncertainties. All four models then estimated further increases in MML under an MSY scenario.

All five models were able to estimate changes in TFB. While they agreed unanimously that TFB has increased during the historic period, this trend was not reflected in the highly uncertain empirical studies. When forecasting to 2050 under a MSY scenario, StrathE2E projected an increase whereas the other four model estimated little additional change.

### 3.2. Uncertainty quantification

The ensemble model predictions for each of the seven indicators are illustrated in Figure 2. MML, TyL and LFI were forecasted to increase and then plateau under an MSY scenario. The uncertainty for these indicators was low throughout and did not grow appreciably over time. In comparison, forecasts for $B M, T F B, Z B$ and $Z: P$ were much more uncertain. In $B M$ and $T F B$ this uncertainty increased into the future, whereas for $Z B$ and $Z: P$ the uncertainty was large throughout. These patterns were consistent across scenarios (see Supplement 6).

### 3.3. Scenario comparison

In general, the MSY and Nash scenarios produced similar results for all seven indicators (see Figures 2 and 3 and Supplement 6, Figure S11) so here we focus solely on MSY scenarios. For MML, TyL, LFI and TFB, the no fishing scenario resulted in the largest absolute indicator values, followed by MSY and then status-quo (Figures 3 and 4). The species composition and size-structure of the fish community should recover under fishing scenarios (MSY and Nash),
with MSY leading to an expectation of a $5 \%, 19 \%$ and $8 \%$ smaller MML, LFI, and TyL respectively, relative to the unfished state (Figure 4). BM was predicted to be higher under the status-quo and no fishing scenarios than it was for MSY. For $Z B$ and $Z: P$, the estimates were associated with large uncertainty independent of the scenario and we are unable to say conclusively which scenario would lead to the largest increases. That said, we estimate that there is a 0.985 and 0.986 probability that the MSY and Nash scenarios will lead to a higher $Z: P$ ratio than under the status quo scenario.

## 4. Discussion

Marine ecosystems face a range of pressures from climate change to the cumulative effects of a growing range of human activities, such as shipping, pollution, or the renewable energy sector (Willsteed et al. 2018). Given our limited understanding on the resilience or vulnerability of ecosystem function, a risk averse strategy would be to minimise human impacts and allow ecosystems to recover to natural levels. However, the socio-economic demands of society also require us to maintain sustainable use of the marine environment. Consequently, biodiversity assessments have been developed to monitor ecological indicators which are designed to track changes in ecosystem state at a range of hierarchical levels (species, communities, and ecosystems). Our study compares strategies for sustainable fishing to demonstrate whether these strategies are likely to lead to decreases in ecological risk. Specifically, this work addresses an explicit policy question that follows from the OSPAR 2017 (OSPAR 2017a b c d, HELCOM 2018) and UK 2019 assessments of biodiversity (HM Government 2019), namely: will fisheries management measures lead to improvements in fish communities (size-structure and species composition) and the biomass of birds and mammals?

The intermediate assessment of OSPAR (2017c d) suggested that the size structure (LFI and $T y L$ ) of demersal fish communities should continue to improve if current fishing levels in the North Sea are maintained and our results confirm that this should be the case in each management scenario. In contrast, empirical estimates by OSPAR (2017b) of MML (measured separately for pelagic and demersal communities) indicate minimal recovery based on survey estimated relative abundance. Here, we modelled the absolute biomass of the whole fish community, not spilt into pelagic and demersal communities, and demonstrate that $M M L$ and TyL should increase in each management scenario. The biomass of birds and mammals should also improve, but we do not expect it to reach the levels seen in the early 1990s. Further, all of the scenarios should lead to improvements in TFB and BM. An aim of EBFM is to support top predators while maintaining productive fisheries (Hill et al. 2020). We show here that fishing on pelagic fish at lower levels than MSY (and Nash), as is the case in status quo and no fishing, will probably lead to greater improvement in $B M$ and $T F B$. This demonstrates a tradeoff between yield and biodiversity, which is important evidence to inform potential management action.

The no-fishing scenario demonstrates the recovery potential of the ecosystem, assuming constant productivity, from the current state and so provides a basis for the quantitative assessment of GES (sensu Rossberg et al. (2017)). The forecast values of each indicator, with their uncertainty, provides an estimate of the 'natural range' of each indicator when the ecosystem is unperturbed. Quantitative targets for indicators can then be made relative to this 'natural range' and this is a particularly useful approach when associated with estimates of temporal responsiveness of change (Houle et al. 2012, Collie et al. 2013, Rossberg et al. 2017). We suggest that this model-based approach is preferable to the simpler empirical approach in which indicators are compared to historic baselines that may no longer be
appropriate (McQuatters-Gollop et al. 2019). Nevertheless, long-term change in abiotic conditions can alter the recovery potential of an ecosystem (for example due to a decrease in ecosystem productivity as suggested by Marshall et al. (2016) and Capuzzo et al. (2018)). Consequently, indicator targets should be iteratively re-evaluated as part of the assessment cycle (Lynam et al. 2016). Similarly, once an indicator target is accepted and management measures to reach the target are considered, then the social and economic effects of the measures should also be considered within the assessment cycle (Lynam et al. 2016).

## Ensemble model

Previous investigations into management effects of ecosystem indicators have used only one ecosystem model per environment and could not account for uncertainty within model projections, potential systematic differences between different modelling approaches or quantify the confidence they have in their prediction (Fu et al. 2019). As a result, the findings of a study can be sensitive to the choice of ecosystem model. As no one model is uniformly better that the rest (Chandler 2013) and indeed all are subject to some degree of structural error, we used a suite of models, combined using the ensemble model of Spence et al. (2018), to generate projections of management effects on ecosystem indicators. This model synthesised information from empirical and modelling studies, exploiting the strengths, while discounting the weaknesses (Chandler 2013), to give probabilistic estimates of the indicators of interest. It integrated multiple sources of uncertainty, including observational, parameter and model uncertainty, to allow a much more robust assessment of risk which is important when giving management advice (Harwood \& Stokes 2003).

In this study we used five ecosystem models to investigate the effect of different management strategies on seven indicators of GES. These ecosystem models were chosen as
they were the only models available to us at the time, and all gave some information about the true state of the ecosystem, both in the past and in the future, under the different fishing scenarios. If more models were available to us (e.g., an Atlantis model, Audzijonyte et al. 2019), there would be more information about the true state of the system and therefore we would have included it in the ensemble study, further reducing the uncertainty of the true value of the indicator and increasing the skill of the suite of models.

In model weighting schemes, e.g., Bayesian model averaging (e.g., Banner \& Higgs 2017, Dormann et al. 2018), or where a single model is used, the uncertainty is limited to the models in the suite, regardless of whether they are good at describing the indicator or not. However, the uncertainty from the ensemble model of Spence et al. (2018) is dependent on the ability of the models to represent the indicators. This is evident when looking at the LFI, where the relative changes in uncertainty for the 2013 and 2050 predictions are small when compared to the TFB, where the relative uncertainty is larger. This is despite only 3 of the models being able to predict $L F I$ but all the models being able to predict $T F B$. The ensemble model provides a way of combining multiple sources of information, both modelling and empirical studies, to predict the outcomes of management scenarios with greater certainty. If managers can be more confident in the outcome of their actions, then they can make better decisions (Zhang 2006).

## Model skill

The uncertainty from the ensemble model can be considered a metric to assess the skill of the suite of models as a collective, with higher uncertainty indicating lower skill. In this study we found that the models were relatively good at predicting the length-based indicators, as the uncertainty remained relatively constant into the future, suggesting higher levels of skill.

When compared to the empirical studies, the three size-based models followed the trends of each length-based indicator well, despite errors in magnitude. Models with predictable errors are often more useful than ones with less. For example, a model that predicts the trends of the system right, despite getting the absolute values wrong, may be more useful, and have greater "model skill", than one that gets the absolute value right, but the trends completely wrong (Spence et al. 2018). Size-based theory has a long history (Sheldon et al. 1972) and the mechanisms that govern growth and size structure are well understood and supported by data (Andersen 2019). The size-based models, mizer, FishSUMs and LeMans, explicitly describe the fish community, with the rest of the ecosystem being implicit, and are all based, to varying degrees, on these mechanisms. On the other hand, EwE and StrathE2E are whole ecosystem models, with the fish community embedded in the wider ecosystem. In these models the processes that act on the size-structure of the fish community are less detailed, suggesting they cannot directly inform the size-based indicators. Instead, they ought to better represent flows of energy in and out of the fish community, which allows one to answer questions about the impact of the whole ecosystem, rather than just the fish community.

For the higher and lower trophic level indicators, the uncertainty is much larger. This is because the models' predictions of zooplankton, phytoplankton, birds and mammals neither reproduced the past adequately, nor agreed with each other when forecasting forwards. In these circumstances, the uncertainty is large, and the forecasts disagree with each other, so the ensemble has lower predictive skill. The lack of skill in the ensemble model for indicators describing the top and bottom of the food web is somewhat understandable. For example, for birds and mammals, their large spatial distribution, migration, and complex behaviour, are less well understood than fish growth (e.g., Newton 2005, Hays et al. 2016). Furthermore, modelling the whole ecosystem, in a computationally tractable way, requires that
simplifications be made and therefore the accuracy of its individual components will be reduced (Skogen et al. 2021). Further development should prioritise the models' abilities to describe the processes governing top and bottom trophic levels. Alternatively, existing models could be coupled (e.g., physics, biogeochemistry and fish habitat models, GarciaGarcia et al. (2021, this issue)), therefore taking advantage of detailed mechanisms at different levels of the ecosystem.

The work presented here uses the ecosystem models designed and built with the data in their respective publications, with the skill reflecting the suite of models as opposed to each individual model. Although the whole ecosystem models do not appear to be as useful as the size-based models in this study, there are many other questions where they may provide useful output, for instance trophic indicators (Bourdaud et al. 2016), that would better demonstrate the utilities of these models. However, in this study we were interested in using multiple sources of information to inform the effects of different fishing strategies on these seven indicators, as opposed to an inter-model comparison, which is beyond the scope of this work. Since this study, there has been further developments of StrathE2E (Heath et al. 2021), which are likely to improve the performance of this ensemble model, particularly for birds and mammals.

## Summary and conclusions

Empirical data analyses currently provide the basis for ecosystem assessment as they represent the key means to monitor change within ecosystems. Data analyses can also demonstrate ecological responses to changes in pressure and the effect of different management measures. However, quantitative modelling tools are the most appropriate basis to explore potential ecological responses to newly planned management measures.

Unfortunately integrating different model outputs is difficult because a) differences in model structure may make it impossible to simply average outputs in a meaningful way, b) the model discrepancies will not be independent of one another, c) uncertainty in the projections can be hard to quantify. The ensemble approach of Spence et al. (2018) directly addresses these issues and provides a powerful framework for integrating a range of models and providing robust advice based on these models and information on their past performance at forecasting quantities of interest, with quantifiable uncertainty (Jardim et al. 2021).

We have used this framework to make projections for seven ecosystem indicators of GES, covering key parts of the North Sea food web from phytoplankton to top predators. All indicators are projected to increase when fishing the community in accordance with MSY principles, and we have high confidence in this recovery for the size-based indicators (LFI, TyL and $M M L$ ). However, our confidence in the future response of $T F B, B M, Z: P$, and $Z B$ is low, with a significant chance that these might not increase with MSY-consistent fishing. Performance with respect to TFB is particularly disappointing as this is a key quantity of interest that is forecast by all five ensemble constituents. Future research should be prioritised to reduce uncertainty in biomass estimates by ecosystem models and empirical analyses because these are important indicators of state which are poorly forecast despite good understanding of size-structure and species composition.

We demonstrated that this new ensemble methodology can be directly applied to address policy-relevant questions and add value for decision-makers because all uncertainties are considered together. From this work we conclude that current fisheries management objectives to achieve MSY can be considered appropriate for aspects of fish communities
(species composition and size structure) as monitored by The Convention for the Protection of the Marine Environment of the North-East Atlantic (OSPAR).

## 5. Acknowledgements

This work was funded by DEFRA via the 'Appraisal of indicators of Good Environmental Status'. MAS and CAG acknowledge funding from the Cefas science development scheme. CAG also acknowledges funding from NERC Highlight grant code NE/T003502/1. We thank Michael Heath and Douglas Speirs for making the StratE2E and FishSUMs models available, Paul Bachem, Brian Wells and three anonymous reviewers for comments on earlier versions of the paper.

## 6. Cited Literature

Andersen KH (2019) Fish Ecology, Evolution, and Exploitation: A New Theoretical Synthesis. Princeton University Press.

Audzijonyte A, Pethybridge H, Porobic J, Gorton R, Kaplan I, Fulton EA (2019) Atlantis: A spatially explicit end-to-end marine ecosystem model with dynamically integrated physics, ecology and socio-economic modules. Methods Ecol Evol 10:1814-1819.

Banner KM, Higgs MD (2017) Considerations for assessing model averaging of regression coefficients. Ecol Appl 27:78-93.

Barneche DR, Ross Robertson D, White CR, Marshall DJ (2018) Fish reproductive-energy output increases disproportionately with body size. Science (80-) 360:642-645.

Bell RJ, Collie JS, Branch TA, Fogarty MJ, Minto C, Ricard D (2018) Changes in the size structure of marine fish communities. ICES J Mar Sci 75:102-112.

Bianchi G, Gislason H, Graham K, Hill L, Jin X, Koranteng K, Manickchand-Heileman S, Payá I, Sainsbury K, Sanchez F, Zwanenburg K (2000) Impact of fishing on size composition and diversity of demersal fish communities. ICES J Mar Sci 57:558-571.

Blanchard JL, Andersen KH, Scott F, Hintzen NT, Piet G, Jennings S (2014) Evaluating targets and trade-offs among fisheries and conservation objectives using a multispecies size spectrum model. J Appl Ecol 51:612-622.

Bourdaud P, Gascuel D, Bentorcha A, Brind'Amour A (2016) New trophic indicators and target values for an ecosystem-based management of fisheries. Ecol Indic 61:588-601.

Brose U, Blanchard JL, Eklöf A, Galiana N, Hartvig M, Hirt MR, Kalinkat G, Nordström MC,

O'gorman EJ, Rall BC, Schneider FD, Thébault E, Jacob U (2017) Predicting the consequences of species loss using size-structured biodiversity approaches. Biol Rev 92:684-697.

Capuzzo E, Lynam CP, Barry J, Stephens D, Forster RM, Greenwood N, McQuatters-Gollop A, Silva T, van Leeuwen SM, Engelhard GH (2018) A decline in primary production in the North Sea over 25 years, associated with reductions in zooplankton abundance and fish stock recruitment. Glob Chang Biol 24:e352-e364.

Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, Brubaker MA, Guo J, Li P, Riddell A (2017) Stan: A probabilistic programming language. J Stat Softw 76.

Casini M, Blenckner T, Möllmann C, Gårdmark A, Lindegren M, Llope M, Kornilovs G, Plikshs M, Stenseth NC (2012) Predator transitory spillover induces trophic cascades in ecological sinks. Proc Natl Acad Sci 109:8185 LP - 8189.

Chandler RE (2013) Exploiting strength, discounting weakness: Combining information from multiple climate simulators. Philos Trans R Soc A Math Phys Eng Sci 371.

Christensen V (1996) Managing fisheries involving predator and prey species. Rev Fish Biol Fish 6:417-442.

Christensen V, Walters CJ (2004) Ecopath with Ecosim: Methods, capabilities and limitations. Ecol Modell 172:109-139.

Christiansen B (2020) Understanding the distribution of multimodel ensembles. J Clim 33:9447-9465.

Coll M, Shannon L, Kleisner KM, Juan-Jordá MJ, Bundy A, Akoglu AG, Banaru D, Boldt JL, Borges MF, Cook A, Diallo I, Fu C, Fox C, Gascuel D, Gurney L, Hattab T, Heymans JJ,

Jouffre D, Knight BR, Kucukavsar S, Large SI, Lynam C, MacHias A, Marshall KN, Masski H, Ojaveer H, Piroddi C, Tam J, Thiao D, Thiaw M, Torres MA, Travers-Trolet M, Tsagarakis K, Tuck I, Van Der Meeren GI, Yemane D, Zador SG, Shin YJ (2016) Ecological indicators to capture the effects of fishing on biodiversity and conservation status of marine ecosystems. Ecol Indic 60:947-962.

Collie J, Rochet M, Bell R (2013) Rebuilding fish communities: the ghost of fisheries past and the virtue of patience. Ecol Appl 23:374-391.

Collie JS, Botsford LW, Hastings A, Kaplan IC, Largier JL, Livingston PA, Plagányi É, Rose KA, Wells BK, Werner FE (2016) Ecosystem models for fisheries management: finding the sweet spot. Fish Fish 17:101-125.

Couce E, Schratzberger M, Engelhard GH (2020) Reconstructing three decades of total international trawling effort in the North Sea. Earth Syst Sci Data 12:373-386.

Dickey-Collas M, Payne MR, Trenkel VM, Nash RDM (2014) Hazard warning: Model misuse ahead. ICES J Mar Sci 71:2300-2306.

Dormann CF, Calabrese JM, Guillera-Arroita G, Matechou E, Bahn V, Bartoń K, Beale CM, Ciuti S, Elith J, Gerstner K, Guelat J, Keil P, Lahoz-Monfort JJ, Pollock L, Reineking B, Roberts DR, Schröder B, Thuiller W, Warton DI, Wintle BA, Wood SN, Wüest RO, Hartig F (2018) Model averaging in ecology: a review of Bayesian, information-theoretic, and tactical approaches for predictive inference. Ecol Monogr 88:485-504.

Eerkes-Medrano D, Fryer RJ, Cook KB, Wright PJ (2017) Are simple environmental indicators of food web dynamics reliable: Exploring the kittiwake-temperature relationship. Ecol Indic 75:36-47.

Engelhard GH, Lynam CP, Bernardo G-C, Dolder PJ, Mackinson S (2015) Effort reduction and the large fish indicator: spatial trends reveal positive impacts of recent European fleet reduction schemes. Environ Conserv 42:227-236.

European Commission (2008) Directive 2008/56/EC of the European Parliament and of the Council Establishing a Framework for Community Action in the Field of Marine Environmental Policy (Marine Strategy Framework Directive). Brussels.

European Union (2015) Regulation (EU) 2015/812 of the European Parliament and of the Council of 20 May 2015 amending Council Regulations (EC) No 850/98, (EC) No 2187/2005, (EC) No 1967/2006, (EC) No 1098/2007, (EC) No 254/2002, (EC) No 2347/2002 and (EC) No 1224/2009, and Regu. Off J Eur Communities L133:1-20.

Fisher JAD, Frank KT, Leggett WC (2010) Global variation in marine fish body size and its role in biodiversity-ecosystem functioning. Mar Ecol Prog Ser 405:1-13.

Froese R, Pauly D (2021) FishBase. World Wide Web Electron Publ.

Fu C, Xu Y, Bundy A, Grüss A, Coll M, Heymans JJ, Fulton EA, Shannon L, Halouani G, Velez L, Akoğlu E, Lynam CP, Shin YJ (2019) Making ecological indicators management ready: Assessing the specificity, sensitivity, and threshold response of ecological indicators. Ecol Indic 105:16-28.

Fulton EA, Smith ADM, Punt AE (2005) Which ecological indicators can robustly detect effects of fishing? ICES J Mar Sci 62:540-551.

Garcia-Garcia LM, Aldridge J, Bruggeman J, Spence MA (2021) A two-way coupled Lower Trophic Level-Higher Trophic level model to investigate the effect of climate change on the ecosystem. Press.

Geary WL, Bode M, Doherty TS, Fulton EA, Nimmo DG, Tulloch AIT, Tulloch VJD, Ritchie EG (2020) A guide to ecosystem models and their environmental applications. Nat Ecol Evol 4:1459-1471.

Gorokhova E, Lehtiniemi M, Postel L, Rubene G, Amid C, Lesutiene J, Uusitalo L, Strake S, Demereckiene N (2016) Indicator properties of baltic zooplankton for classification of environmental status within marine strategy framework directive. PLoS One 11:1-26.

Greenstreet S, Moriarty M (2017) Manual for Version 3 of the Groundfish Survey Monitoring and Assessment Data Product. Scottish Mar Freshw Sci 8:77.

Greenstreet SPR, Fraser HM, Rogers SI, Trenkel VM, Simpson SD, Pinnegar JK (2012) Redundancy in metrics describing the composition, structure, and functioning of the North Sea demersal fish community. ICES J Mar Sci 69:8-22.

Greenstreet SPR, Rogers SI, Rice JC, Piet GJ, Guirey EJ, Fraser HM, Fryer RJ (2011) Development of the EcoQO for the North Sea fish community. ICES J Mar Sci 68:1-11.

Hall SJ, Collie JS, Duplisea DE, Jennings S, Bravington M, Link J (2006) A length-based multispecies model for evaluating community responses to fishing. Can J Fish Aquat Sci 63:1344-1359.

Hartvig M, Andersen KH, Beyer JE (2011) Food web framework for size-structured populations. J Theor Biol 272:113-122.

Harwood J, Stokes K (2003) Coping with uncertainty in ecological advice: lessons from fisheries. Trends Ecol Evol 18:617-622.

Hays GC, Ferreira LC, Sequeira AMM, Meekan MG, Duarte CM, Bailey H, Bailleul F, Bowen WD, Caley MJ, Costa DP, Eguíluz VM, Fossette S, Friedlaender AS, Gales N, Gleiss AC,

Gunn J, Harcourt R, Hazen EL, Heithaus MR, Heupel M, Holland K, Horning M, Jonsen I, Kooyman GL, Lowe CG, Madsen PT, Marsh H, Phillips RA, Righton D, Ropert-Coudert Y, Sato K, Shaffer SA, Simpfendorfer CA, Sims DW, Skomal G, Takahashi A, Trathan PN, Wikelski M, Womble JN, Thums M (2016) Key Questions in Marine Megafauna Movement Ecology. Trends Ecol Evol 31:463-475.

Heath MR, Speirs DC, Steele JH (2014) Understanding patterns and processes in models of trophic cascades. Ecol Lett 17:101-114.

Heath MR, Speirs DC, Thurlbeck I, Wilson RJ (2021) StrathE2E2: An r package for modelling the dynamics of marine food webs and fisheries. Methods Ecol Evol 12:280-287.

HELCOM (2018) State of the Baltic Sea - Second HELCOM Holistic Assessment 2011-2016. Balt Sea Environ Proc 155:155.

Heymans JJ, Coll M, Link JS, 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.

Hiddink JG, Kaiser MJ, Sciberras M, McConnaughey RA, Mazor T, Hilborn R, Collie JS, Pitcher CR, Parma AM, Suuronen P, Rijnsdorp AD, Jennings S (2020) Selection of indicators for assessing and managing the impacts of bottom trawling on seabed habitats. J Appl Ecol 57:1199-1209.

Hill SL, Hinke J, Bertrand S, Fritz L, Furness RW, Ianelli JN, Murphy M, Oliveros-Ramos R, Pichegru L, Sharp R, Stillman RA, Wright PJ, Ratcliffe N (2020) Reference points for predators will progress ecosystem-based management of fisheries. Fish Fish 21:368378.

HM Government (2019) Marine Strategy Part One: UK updated assessment and Good Environmental Status.

HM Government (2010) The Marine Strategy Regulations 2010. Environmental Protection. Marine Management. No. 1627.

HM Government (2020) UK Fisheries Act 2020.

Hoffman MD, Gelman A (2014) The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. J Mach Learn Res 15:1593-1623.

Houle JE, Farnsworth KD, Rossberg AG, Reid DG (2012) Assessing the sensitivity and specificity of fish community indicators to management action. Can J Fish Aquat Sci 69:1065-1079.

Hyder K, Rossberg AG, Allen JI, Austen MC, Barciela RM, Bannister HJ, Blackwell PG, Blanchard JL, Burrows MT, Defriez E, Dorrington T, Edwards KP, Garcia-Carreras B, Heath MR, Hembury DJ, Heymans JJ, Holt J, Houle JE, Jennings S, Mackinson S, Malcolm SJ, McPike R, Mee L, Mills DK, Montgomery C, Pearson D, Pinnegar JK, Pollicino M, Popova EE, Rae L, Rogers SI, Speirs D, Spence MA, Thorpe R, Turner RK, van der Molen J, Yool A, Paterson DM (2015) Making modelling count - increasing the contribution of shelf-seas community and ecosystem models to policy development and management. Mar Policy 61:291-302.

ICES (2021) Advice on fishing opportunities. In: Report of the ICES Advisory Committee, 2021. ICES Advice 2021, section 1.1.1.

ICES (2009) Report of the Working Group on the Ecosystem Effects of Fishing Activities (WGECO). ICES C 2009/ACOM20.

ICES (2014) Report of the Working Group on the Ecosystem Effects of Fishing Activities (WGECO). ICES Doc C 2014/ACOM26, Sect 343.

ICES (2016) Working Group on Multispecies Assessment Methods (WGSAM). ICES Sci Reports 1:91:320 pp.

Jardim E, Azevedo M, Brodziak J, Brooks EN, Johnson KF, Klibansky N, Millar CP, Minto C, Mosqueira I, Nash RDM, Vasilakopoulos P, Wells BK (2021) Operationalizing ensemble models for scientific advice to fisheries management. ICES J Mar Sci:fsab010.

Jennings S, Kaiser MJ (1998) The effects of fishing on marine ecosystems. Adv Mar Biol:201212.

Landi P, Minoarivelo HO, Brännström Å, Hui C, Dieckmann U (2018) Complexity and stability of ecological networks: a review of the theory. Popul Ecol 60:319-345.

Lynam CP, Halliday NC, Ho H, Wright PJ, Damme CJG Van, Edwards M, Pitois SG (2013) Spatial patterns and trends in abundance of larval sandeels in the North Sea: 19502005. ICES J Mar Sci 70:540-553.

Lynam CP, Llope M, Möllmann C, Helaouët P, Bayliss-Brown GA, Stenseth NC (2017) Interaction between top-down and bottom-up control in marine food webs. Proc Natl Acad Sci 114:1952 LP - 1957.

Lynam CP, Uusitalo L, Patrício J, Piroddi C, Queirós AM, Teixeira H, Rossberg AG,
Sagarminaga Y, Hyder K, Niquil N, Möllmann C, Wilson C, Chust G, Galparsoro I, Forster R, Veríssimo H, Tedesco L, Revilla M, Neville S (2016) Uses of innovative modeling tools within the implementation of the marine strategy framework directive. Front Mar Sci 3:1-18.

Mackinson S, Platts M, Garcia C, Lynam C (2018) Evaluating the fishery and ecological consequences of the proposed North Sea multi-annual plan. PLoS One 13:1-23.

Marshall AM, Bigg GR, van Leeuwen SM, Pinnegar JK, Wei HL, Webb TJ, Blanchard JL (2016) Quantifying heterogeneous responses of fish community size structure using novel combined statistical techniques. Glob Chang Biol 22:1755-1768.

McQuatters-Gollop A, Mitchell I, Vina-Herbon C, Bedford J, Addison PFE, Lynam CP, Geetha PN, Vermeulan EA, Smit K, Bayley DTI, Morris-Webb E, Niner HJ, Otto SA (2019) From science to evidence - how biodiversity indicators can be used for effective marine conservation policy and management. Front Mar Sci 6:1-16.

Moriarty M, Greenstreet SPR, Rasmussen J (2017) Derivation of Groundfish Survey Monitoring and Assessment Data Product for the Northeast Atlantic Area. Scottish Mar Freshw Sci 8:240.

Newton I (2005) New insights on the ecology of bird migration. Trends Ecol Evol 20:292293.

Olsen E, Fay G, Gaichas S, Gamble R, Lucey S, Link JS (2016) Ecosystem model skill assessment. Yes We Can! PLoS One 11:1-24.

OSPAR (2017a) Changes in Phytoplankton and Zooplankton Communities. Intermed Assess 2017.

OSPAR (2009) Evaluation of the OSPAR system of Ecological Quality Objectives for the North Sea (update 2010).

OSPAR (2017b) Pilot Assessment of Mean Maximum Length of Fish. Intermed Assess 2017.

OSPAR (2017c) Proportion of Large Fish (Large Fish Index). Intermed Assess 2017.

OSPAR (2017d) Size Composition in Fish Communities. Intermed Assess 2017.

Palomares MLD, Froese R, Derrick B, Meeuwig JJ, Nöel SL, Tsui G, Woroniak J, Zeller D, Pauly D (2020) Fishery biomass trends of exploited fish populations in marine ecoregions, climatic zones and ocean basins. Estuar Coast Shelf Sci 243.

Pauly D, Christensen V, Dalsgaard J, Froese R, Torres F (1998) Fishing down marine food webs. Science (80- ) 279:860-863.

Pitois SG, Fox CJ (2006) Long-term changes in zooplankton biomass concentration and mean size over the Northwest European shelf inferred from Continuous Plankton Recorder data. ICES J Mar Sci 63:785-798.

Plagányi ÉE (2007) Models for an ecosystem approach to fisheries. Rome, FAO.

Polovina JJ (1984) Model of a coral reef ecosystem - I. The ECOPATH model and its application to French Frigate Shoals. Coral Reefs 3:1-11.

Queirós AM, Fernandes J, Genevier L, Lynam CP (2018) Climate change alters fish community size-structure, requiring adaptive policy targets. Fish Fish 19:613-621.

R Core Team (2020) A Language and Environment for Statistical Computing. R Found Stat Comput.

Rochet MJ, Collie JS, Jennings S, Hall SJ (2011) Does selective fishing conserve community biodiversity? Predictions from a length-based multispecies model. Can J Fish Aquat Sci 68:469-486.

Rosenberg AA, Kleisner KM, Afflerbach J, Anderson SC, Dickey-Collas M, Cooper AB, Fogarty

MJ, Fulton EA, Gutiérrez NL, Hyde KJW, Jardim E, Jensen OP, Kristiansen T, Longo C, Minte-Vera C V, Minto C, Mosqueira I, Osio GC, Ovando D, Selig ER, Thorson JT, Walsh JC, Ye Y (2018) Applying a New Ensemble Approach to Estimating Stock Status of Marine Fisheries around the World. Conserv Lett 11:e12363.

Rossberg AG, Gaedke U, Kratina P (2019) Dome patterns in pelagic size spectra reveal strong trophic cascades. Nat Commun 10:1-11.

Rossberg AG, Uusitalo L, Berg T, Zaiko A, Chenuil A, Uyarra MC, Borja A, Lynam CP (2017) Quantitative criteria for choosing targets and indicators for sustainable use of ecosystems. Ecol Indic 72:215-224.

Rougier J (2007) Probabilistic Inference for Future Climate Using an Ensemble of Climate Model Evaluations. Clim Change 81:247-264.

Scott F, Blanchard JL, Andersen KH (2014) Mizer: An R package for multispecies, trait-based and community size spectrum ecological modelling. Methods Ecol Evol 5:1121-1125.

Sheldon RW, Prakash A, Sutcliffe Jr. WH (1972) The size distribution of particles in the ocean. Limnol Oceanogr 17:327-340.

Sherley RB, Ladd-Jones H, Garthe S, Stevenson O, Votier SC (2020) Scavenger communities and fisheries waste: North Sea discards support 3 million seabirds, 2 million fewer than in 1990. Fish Fish 21:132-145.

Skogen MD, Ji R, Akimova A, Daewel U, Hansen C, Hjøllo SS, van Leeuwen SM, Maar M, Macias D, Mousing EA, Almroth-Rosell E, Sailley SF, Spence MA, Troost TA, van de Wolfshaar K (2021) Disclosing the truth: are models better than observations? Mar Ecol Prog Ser Press.

Speirs DC, Greenstreet SPR, Heath MR (2016) Modelling the effects of fishing on the North Sea fish community size composition. Ecol Modell 321:35-45.

Speirs DC, Guirey EJ, Gurney WSC, Heath MR (2010) A length-structured partial ecosystem model for cod in the North Sea. Fish Res 106:474-494.

Spence MA, Bannister HJ, Ball JE, Dolder PJ, Griffiths CA, Thorpe RB (2020) LeMaRns: A Length-based Multi-species analysis by numerical simulation in R. PLoS One 15:e0227767.

Spence MA, Blackwell PG, Blanchard JL (2016) Parameter uncertainty of a dynamic multispecies size spectrum model. Can J Fish Aquat Sci 73:589-597.

Spence MA, Blanchard JL, Rossberg AG, Heath MR, Heymans JJ, Mackinson S, Serpetti N, Speirs DC, Thorpe RB, Blackwell PG (2018) A general framework for combining ecosystem models. Fish Fish 19:1031-1042.

Spence MA, Thorpe RB, Blackwell PG, Scott F, Southwell R, Blanchard JL (2021) Quantifying uncertainty and dynamical changes in multi-species fishing mortality rates, catches and biomass by combining state-space and size-based multi-species models. Fish Fish 00:115.

Tam JC, Link JS, Rossberg AG, Rogers SI, Levin PS, Rochet M-J, Bundy A, Belgrano A, Libralato S, Tomczak M, van de Wolfshaar K, Pranovi F, Gorokhova E, Large SI, Niquil N, Greenstreet SPR, Druon J-N, Lesutiene J, Johansen M, Preciado I, Patricio J, Palialexis A, Tett P, Johansen GO, Houle J, Rindorf A (2017) Towards ecosystem-based management: identifying operational food-web indicators for marine ecosystems. ICES J Mar Sci 74:2040-2052.

Thorpe RB, Jennings S, Dolder PJ (2017) Risks and benefits of catching pretty good yield in multispecies mixed fisheries. ICES J Mar Sci 74:2097-2106.

Thorpe RB, Le Quesne WJF, Luxford F, Collie JS, Jennings S (2015) Evaluation and management implications of uncertainty in a multispecies size-structured model of population and community responses to fishing. Methods Ecol Evol 6:49-58.

Tittensor DP, Eddy TD, Lotze HK, Galbraith ED, Cheung W, Barange M, Blanchard JL, Bopp L, Bryndum-Buchholz A, Büchner M, Bulman C, Carozza DA, Christensen V, Coll M, Dunne JP, Fernandes JA, Fulton EA, Hobday AJ, Huber V, Jennings S, Jones M, Lehodey P, Link JS, Mackinson S, Maury O, Niiranen S, Oliveros-Ramos R, Roy T, Schewe J, Shin Y-J, Stock CA, Underwood PJ, Volkholz J, Watson JR, Walker ND (2017) A protocol for the intercomparison of marine fishery and ecosystem models: Fish-MIP v1.0. Geosci Model Dev Discuss:1-39.

Tu CY, Chen KT , Hsieh CH (2018) Fishing and temperature effects on the size structure of exploited fish stocks. Sci Rep 8:1-10.

UKMMAS (2019) UK Marine Online Assessment Tool.

Waggitt JJ, Evans PGH, Andrade J, Banks AN, Boisseau O, Bolton M, Bradbury G, Brereton T, Camphuysen CJ, Durinck J, Felce T, Fijn RC, Garcia-Baron I, Garthe S, Geelhoed SCV, Gilles A, Goodall M, Haelters J, Hamilton S, Hartny-Mills L, Hodgins N, James K, Jessopp M, Kavanagh AS, Leopold M, Lohrengel K, Louzao M, Markones N, Martínez-Cedeira J, Ó Cadhla O, Perry SL, Pierce GJ, Ridoux V, Robinson KP, Santos MB, Saavedra C, Skov H, Stienen EWM, Sveegaard S, Thompson P, Vanermen N, Wall D, Webb A, Wilson J, Wanless S, Hiddink JG (2020) Distribution maps of cetacean and seabird populations in
the North-East Atlantic. J Appl Ecol 57:253-269.

Walker ND, Maxwell DL, Le Quesne WJF, Jennings S (2017) Estimating efficiency of survey and commercial trawl gears from comparisons of catch-ratios. ICES J Mar Sci 74:14481457.

Willsteed EA, Birchenough SNR, Gill AB, Jude S (2018) Structuring cumulative effects assessments to support regional and local marine management and planning obligations. Mar Policy 98:23-32.

Zhang FX (2006) Information Uncertainty and Stock Returns. J Finance 61:105-137. ordered by asymptotic size.

| Species (Latin name) | $\boldsymbol{F}_{\text {MSY }}$ | $\boldsymbol{F}_{\text {Nash }}$ | $\boldsymbol{F}_{\text {Status }}$ quo |
| :--- | :---: | :---: | :---: |
| Sprat (Sprattus sprattus) | 1.30 | 0.78 | 1.06 |
| Norway pout (Trisopterus esmarkii) | 0.35 | 0.63 | 0.48 |
| Sandeel (Ammodytes tobianus) | 0.35 | 0.63 | 0.57 |
| Poor cod (Trisopterus minutus) | 0.72 | 0.58 | 0.55 |
| Long rough dab (Hippoglossoides | 0.60 | 0.54 | 0.55 |
| platessoides) | 0.41 | 0.57 | 0.55 |
| Dab (Limanda limanda) | 0.25 | 0.50 | 0.25 |
| Herring (Clupea harengus) | 0.50 | 0.55 | 0.63 |
| Horse mackerel (Trachurus trachurus) | 0.33 | 0.48 | 0.55 |
| Lemon sole (Microstomus kitt) | 0.22 | 0.42 | 0.53 |
| Sole (Solea solea) | 0.32 | 0.45 | 0.63 |
| Mackerel (Scomber scombrus) | 0.21 | 0.42 | 0.37 |
| Whiting (Merlangius merlangus) | 0.27 | 0.38 | 0.55 |
| Witch (Glyptocephalus cynoglossus) | 0.27 | 0.46 | 0.55 |
| Grey gurnard (Eutrigla gurnardus) | 0.25 | 0.20 | 0.55 |
| Plaice (Pleuronectes platessa) | 0.15 | 0.32 | 0.55 |
| Starry ray (Amblyraja radiata) | 0.30 | 0.42 | 0.58 |
| Haddock (Melanogrammus aeglefinus) | 0.11 | 0.14 | 0.55 |
| Cuckoo ray (Leucoraja naevus) | 0.10 | 0.23 | 0.55 |
| Monkfish (Lophius piscatorius) | 0.21 | 0.88 |  |
| Cod (Gadus morhua) | 0.30 | 0.38 |  |
| Saithe (Pollachius virens) |  |  |  |

## Tables

Table 1. Species list and species-specific fishing mortality rates for MSY and Nash scenarios taken from Thorpe et al. (2017), with some species having a decreased fishing mortality in MSY and Nash, relative to status quo and others having an increased mortality. Species are


2 Figure 1. Indicator estimates from the combination of all five ecosystem models and the empirical studies until 2013. Starting in 2014, each ecosystem model is run under the MSY fishing scenario. The uncertainty envelope represents $\pm 2$ standard deviations. The black vertical line indicates the last year for which empirical studies were considered.


Figure 2. Indicator predictions from the ensemble model under the MSY fishing scenario. The uncertainty envelope surrounding each prediction represents $\pm 2$ standard deviations. Note,

3 the empirical studies end in 2013 as illustrated by the black vertical line.


Figure 3: Violin plots of the ensemble model's predictions for the indicators in 2013 and 2050 under the different management scenarios (SQ = Status Quo, MSY = Maximum Sustainable Yield, Nash = Nash equilibrium, NF = No fishing).


LFI



MML
 here.
Figure 4. Difference between management scenarios (NF = No fishing, SQ = Status Quo, MSY $=$ Maximum Sustainable Yield) in the long-term predictions for each indicator (see separate pane titles) from the ensemble model. The pane in the top left-hand corner demonstrates how to read the plots i.e., areas where one scenario leads to a better solution than another. The circles encapsulate the uncertainty in the outcomes showing the $25^{\text {th }}, 75^{\text {th }}$ and $95^{\text {th }}$ percentiles. The MSY and Nash scenarios lead to very similar results and so only MSY is shown


Figure 4. Continued.


[^0]:    Take down policy
    If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

[^1]:    Sustainable fishing can lead to improvements in marine ecosystem status: an ensemblemodel forecast of the North Sea ecosystem

    Michael A. Spence ${ }^{1,{ }^{*}}$, Christopher A. Griffiths ${ }^{1,2}$, James J. Waggitt ${ }^{3}$, Hayley J. Bannister ${ }^{1}$, Robert B. Thorpe ${ }^{1}$, Axel G. Rossberg ${ }^{4}$, Christopher P. Lynam ${ }^{1}$
    ${ }^{1}$ Centre for Environment, Fisheries and Aquaculture Science, Lowestoft Laboratory, Lowestoft, Suffolk NR33 OHT, UK.
    ${ }^{2}$ Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield, S10 2TN, UK.
    ${ }^{3}$ School of Ocean Sciences, Bangor University, Menai Bridge, LL59 5AB, UK.
    ${ }^{4}$ School of Biological and Chemical Sciences, Queen Mary University of London, London, E1 4NS, UK.
    *Corresponding author: Email - michael.spence@cefas.co.uk

    Running page head: Ensemble-modelling of marine indicators

