## Supplement 1 – Statistical estimation of zooplankton biomass (ZB) and biomass of birds and mammals (BM)

From the CPR dataset, samples for 29 taxa (see Table S1) were modelled and each sample was assumed to represent a volume filtered of 3.11 m$^3$ (John et al. 2002). These taxa were selected since factors were available for them to correct for the known under-sampling by the CPR (that samples at a fixed depth ~10 m) relative to the WP-2 plankton net that samples the whole water column with 100% filtration efficiency. Abundance (individuals) to dry weight (mg) conversion factors were also given from Pitois & Fox (2006). Once sample estimates were corrected for under-sampling (including haul depth) and converted into dry weight, these estimates were averaged by month for each euphotic zone (EHZ) and assumed to broadly represent the biomass per unit area (mg m$^{-2}$) in the EHZ (Figure S1). We then model the seasonality and long-term change using a Generalised Additive Model (GAM) applied to each of the zones where the response variable is the log10(1 + mg per m$^2$) and the explanatory variable is a 2D tensor smooth applied to Year and Month, assuming in normal link function. The output, once back-transformed, from each of these models for each month was raised up to the geographic area of the EHZ and summed to give the total annual biomass per EHZ. The total estimated biomass for the North Sea was calculated from the sum of these EHZ estimates and scaled up to wet weight (using the conversion factor of 5, Pitois & Fox (2006)) and scaled again to represent the whole community (not only the selected species, unfortunately the only way to do this was by using the proportion of abundance).

The GAMs fitted significantly to each subset of data (Figure S1) with a high proportion of deviance explained by the models (p < 0.0001, Table S2) and well-behaved residuals (Figure S2). The biomass in each EHZ other than PMX showed an increase in biomass during the 1980s followed by a fall in the 1990s, the perm mixed zone decreased throughout the time-series (Figure S3). The total biomass of the CPR sampled zooplankton community in the North Sea was found to vary between 10 and 16 million tonnes (Figure S4).

Birds and mammals (*BM)* was taken from species distribution model (SDM) predictions based on ~2 million kilometres of at-sea cetacean and seabird surveys between 1985 and 2015 (Waggitt et al. 2020). SDM use associations between animals and environmental conditions to predict densities of the former in space and time (Elith & Leathwick 2009). As the existing SDM only produced estimates of BM for a typical calendar month between 1986 and 2018, modifications were needed to provide estimates of BM for a specific calendar month. To provide these estimates, 5 year moving averages of the North Atlantic Oscillation (NAO) and Atlantic Multidecadal Oscillation (AMO) index were added to the existing suite of environmental conditions in the SDM. Both NAO and AMO are known to influence cetacean and seabird populations in the North Sea (Evans & Waggitt 2020, Mitchell et al. 2020). These environmental conditions were added to the presence-absence component for cetaceans, and the count component for seabirds (see Waggitt et al. (2020) for further details). These differences between taxa were because seabird movement is constrained by their need to return to terrestrial colonies for large parts of the year. Therefore, impacts of AMO/NAO on seabirds most likely lower reproductive rates and survival, causing declines in numbers of animals (Mitchell et al. 2020). By contrast, cetacean movement is not constrained. As a consequence, impacts of AMO/NAO are likely distribution shifts, causing changes in encounter rates (Evans & Waggitt 2020).

Table S1. Taxa selected for biomass estimation.

| Group | Taxa | Dry weight (mg) | Correction factor (mean WP-2/CPR ratio) |
|---|---|---|---|
| G1 | *Evadne* spp. | 0.0018 | 18.2 |
| | *Oithona* spp. | 0.0022 | 45.4 |
| | *Penilia* spp. | 0.0034 | 20.1 |
| | *Podon* spp. | 0.0036 | 5.1 |
| | *Para-Pseudocalanus* spp. | 0.0042 | 13.9 |
| G2 | *Corycaeus* spp. | 0.062 | 26.8 |
| | *Acartia* spp. | 0.0082 | 8.7 |
| | *Clausocalanus* spp. | 0.0106 | 10.3 |
| | *Pseudocalanus elongatus* | 0.0109 | 11.6 |
| | *Metridia* I-IV | 0.0124 | 20 |
| | *Temora longicornis* | 0.0126 | 18.3 |
| G3 | *Calanus* I-IV | 0.013 | 2.4 |
| | *Isias clavipe* | 0.0134 | 11.6 |
| | *Centropages hamatus* | 0.0147 | 1.7 |
| | *Candacia* I-IV | 0.0237 | 8.8 |
| | *Centropages typicus* | 0.036 | 1.7 |
| | *Calanoides carinatus* | 0.0533 | 5.7 |
| | *Metridia lucens* | 0.0623 | 20 |
| | *Euchaeta hebes* | 0.0798 | 4.5 |
| | *Candacia armata* | 0.1062 | 3.8 |
| | *Calanus helgolandicus* | 0.1096 | 2.4 |
| G4 | *Pleuromamma robusta* | 0.1877 | 2.7 |
| | *Euchaeta acuta* | 0.2372 | 2.4 |
| | *Euchirella rostrata* | 0.3045 | 2.1 |
| | *Eucalanus crassus* | 0.3055 | 2.1 |
| | *Calanus finmarchicus* | 0.3192 | 2.4 |
| | *Undeuchaeta plumosa* | 0.3281 | 2.1 |
| G5 | *Euchaeta norvegica* | 2.8411 | 2.4 |
| | *Calanus hyperboreus* | 2.85 | 2.4 |

Table S2. Percentage of deviance explained by each GAM and geographic area of the ecohydrodynamic unit (EHU).

| EHU | % of deviance | Km$^2$ |
|---|---|---|
| Seasonally stratified | 72.4 | 174800 |
| Permanently mixed | 59.3 | 49100 |
| Freshwater influence | 44.0 | 22670 |
| Intermittently stratified | 56.7 | 60960 |
| Transitional East | 53.7 | 78050 |
| Transitional West | 67.5 | 30790 |
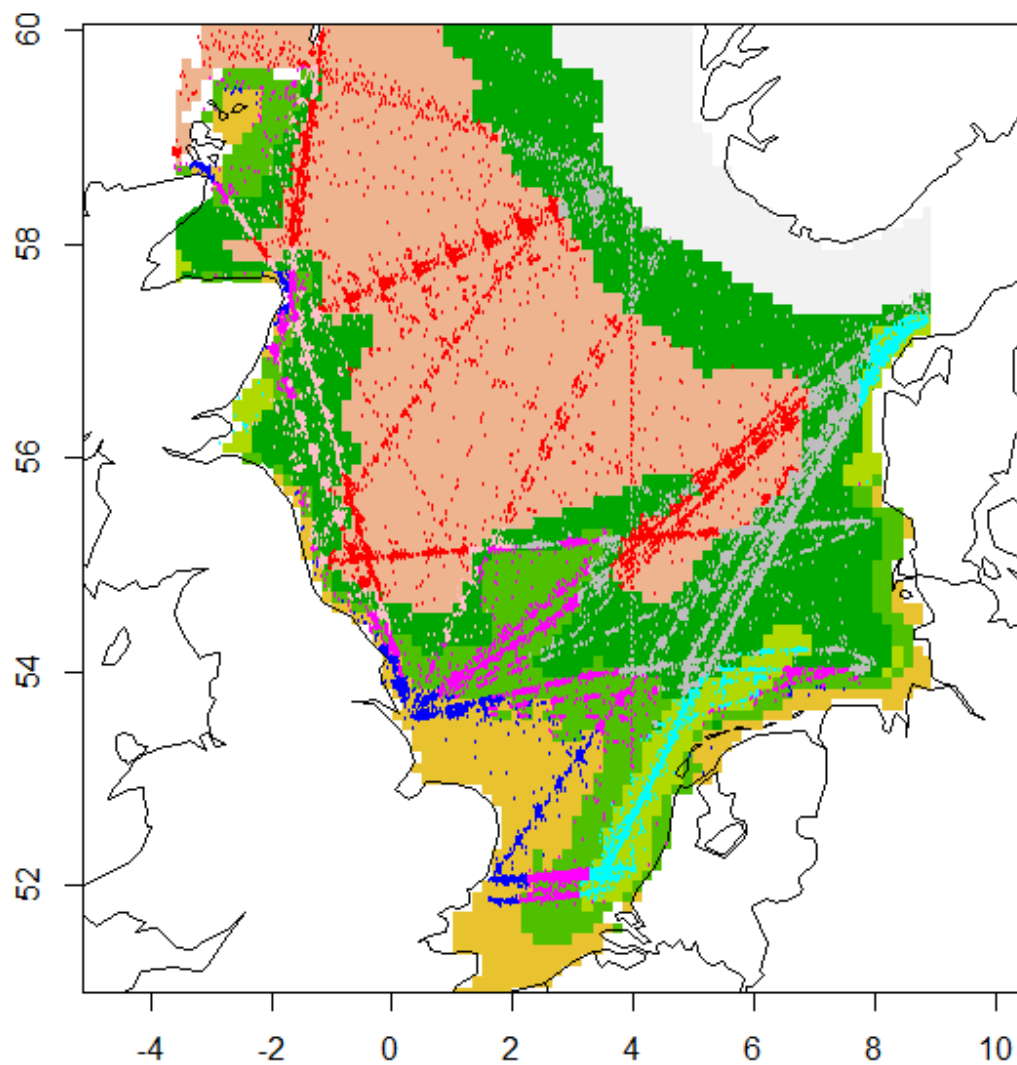
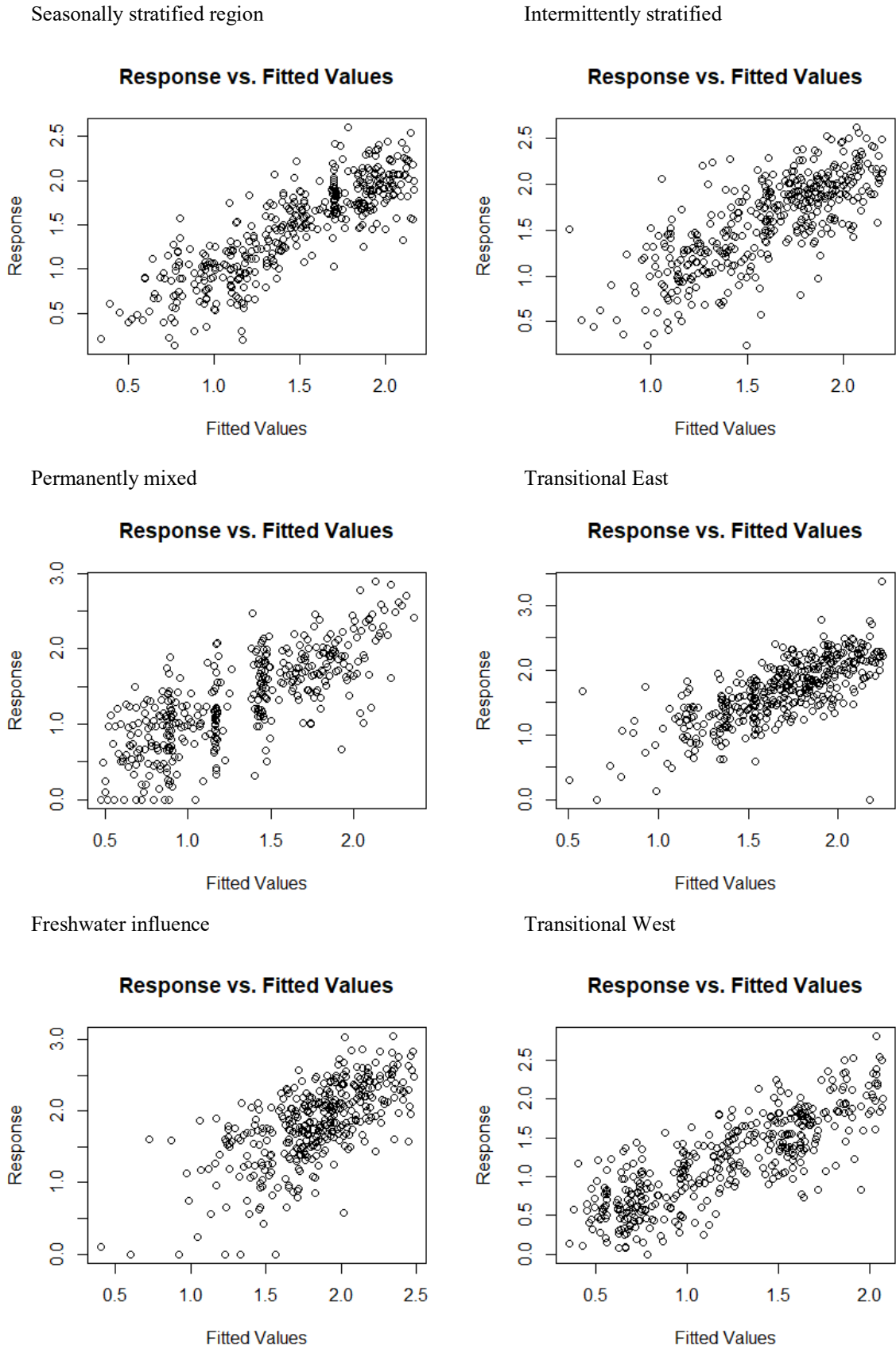Figure S1. CPR sampling locations coloured by ecohydrodynamic unit (EHU).

Seasonally stratified region

Intermittently stratified

**Response vs. Fitted Values**

**Response vs. Fitted Values**

Permanently mixed

Transitional East

**Response vs. Fitted Values**

**Response vs. Fitted Values**

Freshwater influence

Transitional West

**Response vs. Fitted Values**

**Response vs. Fitted Values**

Figure S2. Linearity between fitted values and response for each GAM.
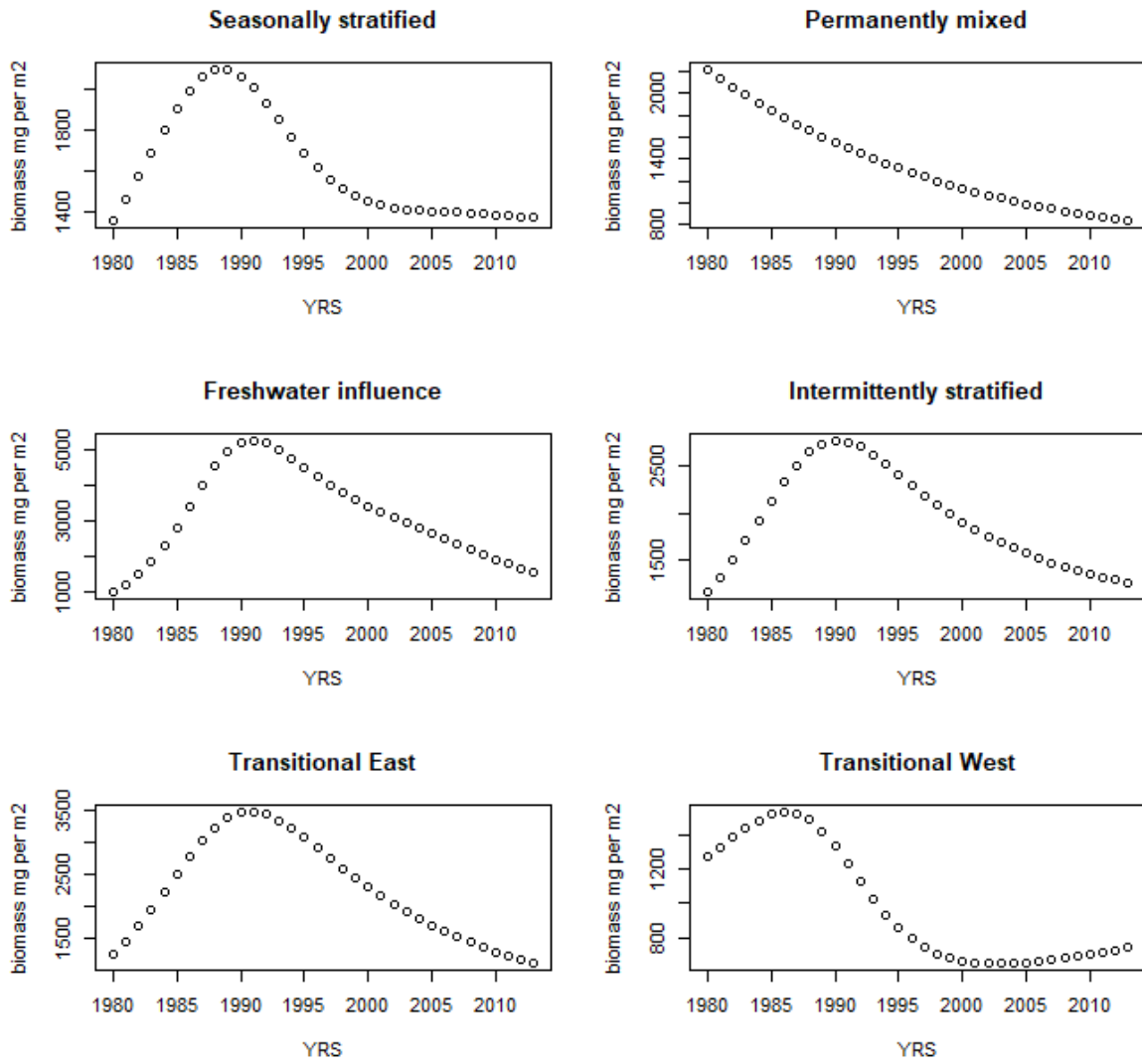
Figure S3. Estimated biomass (dry weight, mg per m$^2$) in each euphotic zone (EHZ).
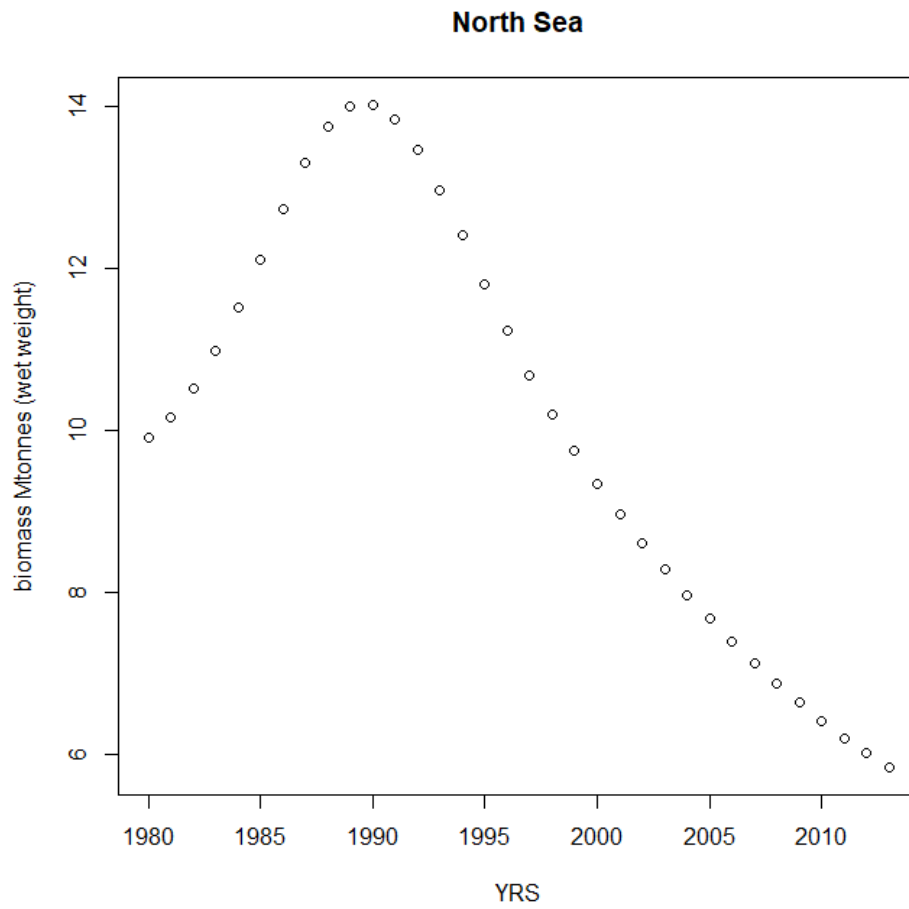
**North Sea**

Figure S4. Estimated biomass (wet weight, million tonnes) in the North Sea.

## Supplement 2 – Polynomial regression of fish observations

Observations of the size structure and size composition indicators (*MML*, *TyL* and *LFI*), as well observations of total fish biomass (*TFB*) were subjected to a 2nd order multivariate polynomial regression, i.e.

$$f(\boldsymbol{y}_t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \boldsymbol{\epsilon}_t, \quad (S1)$$

where $t$ is the year,

$$\boldsymbol{\epsilon}_t \sim MVN(\boldsymbol{0}, \Sigma), \quad (S2)$$

and

$$f(\boldsymbol{y}_t) = \begin{pmatrix} ln(MML) \\ ln(TYL) \\ logit(LFI) \\ ln(TYL) \end{pmatrix} \quad (S3)$$

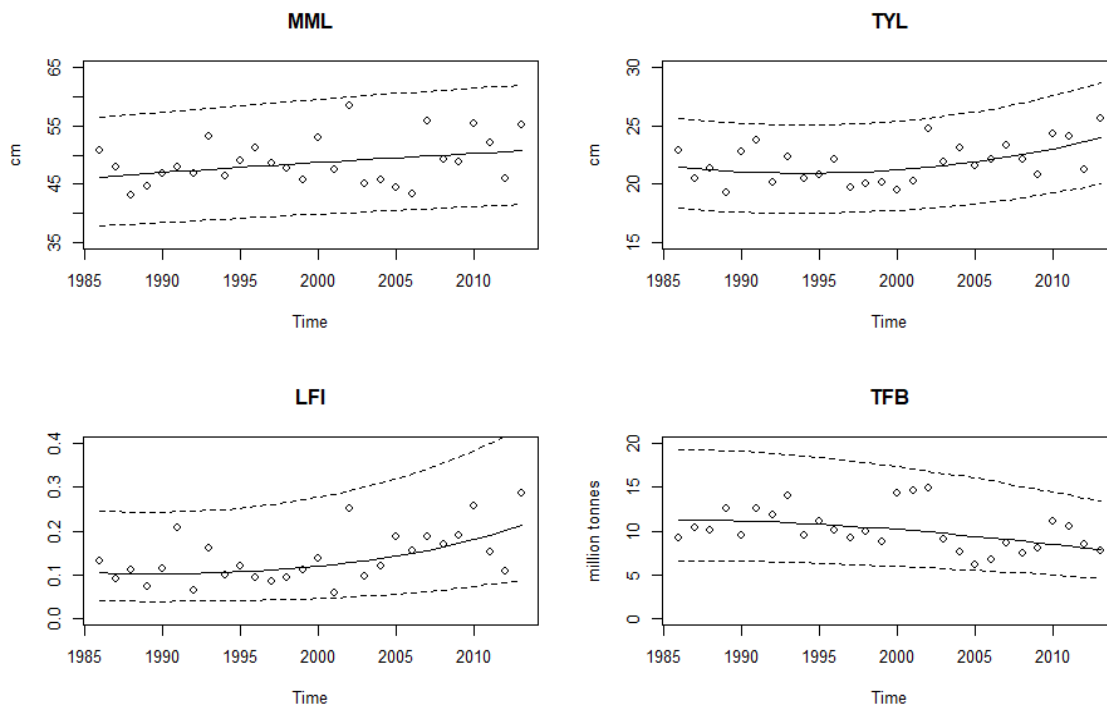The results of this smoothing process are shown in Figure S5.



Figure S5. Empirical observations fitted with a second order polynomial regression.

## Supplement 3 – Scenarios

In this section we describe the fishing scenarios that we examined.

*Status quo*

In this scenario we fish all species at the levels they were fished in 2013.

*Closure*

In this scenario we stop fishing all species in each of the respect models 2013.

*Maximum* sustainable yield

Fish stocks are often managed by considering the fishing mortality that maximises the long-term yield, i.e., the Maximum Sustainable Yield (MSY). We can define $f_i(F_i, \boldsymbol{F}_{-i})$ as the $i$th stock's long-term yield, where $F_i$ is the fishing mortality of the $i$th stock and $\boldsymbol{F}_{-i}$ are the fishing mortalities of the other stocks. Many stocks are managed on a stock-by-stock basis using single-species models. This means that

$$f_i(F_i, \boldsymbol{F}_{-i}) = f_i(F_i, \boldsymbol{F'}_{-i}) \quad \text{(S4)}$$

$\forall \boldsymbol{F'}_{-i}$, and then

$$F_{MSY,i} = \text{argmax}_{F_i} f_i(F_i, \boldsymbol{F}_{-i}) \quad \text{(S5)}$$

is commonly well defined. The fishing mortality rates for the maximum sustainable yield scenario used in this study are from Thorpe et al. (2017) and are shown in Table 1 (see manuscript). All other species were fished at 2013 levels.

*Nash* equilibrium

Stocks often interact with one another and the fishing mortality of the $j$th stock affects the catch of the $i$th stock, i.e.

$$f_i(F_i, \boldsymbol{F}_{-i}) = f_i(F_i, \boldsymbol{F'}_{-i}) \quad \text{(S6)}$$

We therefore need to define a multi-species MSY. One possibility is the Nash equilibrium, which is defined as the point at which we are unable to increase $f_i(F_i, \boldsymbol{F}_{-i})$ by changing $F_i$ only, $\forall i$. Formally, $F_{Nash,i}$ is a Nash equilibrium when

$$\forall i, F_i: f_i(F_{Nash,i}, \boldsymbol{F}_{Nash,-i}) \geq f_i(F_i, \boldsymbol{F}_{Nash,-i}) \quad \text{(S7)}$$

The fishing mortality rates for the Nash scenario used in this study are from Thorpe et al. (2017) and are shown in Table 1 (see manuscript). All other species were fished at 2013 levels.

## Supplement 4 – Model summaries

### Mizer

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. In this form, the model has principally been used to describe the effects of fishing on interacting species and the size-spectrum. Mizer provides predictions of the abundance of each species at 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). It thus differs from some other size-based models that assume deterministic growth based on life history parameters. The smallest individuals in the model do not eat fish belonging to the fish populations but consume smaller planktonic or benthic organisms which we describe as a background resource spectrum. Fish grow and die according to size-dependent predation and, if mature, recruit new young which are put back into the system at the minimum weight. The model can predict abundance at size, biomass, growth and mortality rates for each species. For a complete description of the model see Hartvig et al. (2011) or Scott et al. (2014). Blanchard et al. (2014) developed and applied a version of mizer for the North Sea. In the model, 12 of the more common species have been explicitly represented (sandeel, Norway pout, herring, whiting, grey gurnard, Dab, sole, plaice, haddock, cod, saithe). It is this version of mizer that has been used in this study. We used parameter values from Spence et al. (2016) to simulate up until 2010 and, assuming conditionally independent Gaussian errors on the landings, we used a particle filter (see Doucet & Johansen (2008)) for an introduction to particle filters) to update the fishing mortalities for 2011-13. 100 samples from the joint posterior distribution were simulated from 1968-2100 with fishing being turned off in 2013.

### Ecopath with Ecosim

Ecopath was developed first in 1984 by Polovina (1984) and has been updated subsequently to include temporal (Ecosim) and spatial (Ecospace) dynamics (Christensen & Walters 2004) and is currently used extensively to simulate historic changes in ecosystems (Heymans et al. 2016). The Ecopath model used in the case study is the model of the North Sea (Lynam & Mackinson 2015). It contains > 10 fishing fleets and > 60 functional groups and some of which are split into multiple age stanzas. We used parameter uncertainty from Mackinson et al. (2018).

### StrathE2E

The Strathclyde end-to-end (StrathE2E) marine food web model was designed to simulate regional scale, macroscopic top-down and bottom-up cascading trophic effects (Heath et al. 2014b). 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. The scheme takes off-line output from General Circulation Models (GCM) in the form of volume and nutrient fluxes through the external boundaries and mixing rates between the vertical compartments but is not directly coupled to any GCM. The advantage is

very fast run-times which has enabled the implementation of computational parameter optimisation methods to fit the models to observed data, sensitivity analysis (Morris et al. 2014), and computation of likelihoods for model outputs. The focus of existing uses of Strathe2E has been on UK shelf seas and the cascading implications of fisheries and fishing practices such as trawling and its impacts on the seabed and discarding of unwanted catch (Heath 2012, Heath et al. 2014a, 2015).

## FishSUMS

The FishSUMS model (Speirs et al. 2010, 2016) represents the population dynamics of a set of key trophically-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 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". Outputs from the model are time series of total species biomass (TSB), normalised length distributions at annual census dates, annual recruitment, catch and landings, for each of the focal species. The model was initially configured for the North Sea with a set of nine structured species focused on cod and its main predators and prey (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). In general, the model is configurable for any set of structured species and unstructured prey groups. The model has been developed as a package for the R software environment. The North Sea model includes: sprat, sandeel, Norway pout, herring, whiting, grey gurnard, plaice, haddock, cod, saithe and nephrops.

## LeMans

The LeMans North Sea model framework (Thorpe et al. 2015, 2016, 2017) is an ensemble of length-structured multispecies models which account for multispecies interactions and model parameter uncertainty. It is a modified form of the length-based multispecies model initially developed by Hall et al. (2006) to represent the Georges Bank fish community, and 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 5cm each, spanning the full-size range of species represented into the model (nearly 200cm for some simulations). Progression of individuals through length classes is represented by a deterministic von Bertalanffy growth equation. Individuals mature when they reach a certain size which is defined by a logistic model, in which 50% of the individuals maturing at the length of maturity ($L_{mat}$ see Table S2 in Thorpe et al. (2017)). Reproduction is described with a spawner recruit relationship, which determines the numbers of recruits entering the smallest size class from the biomass of mature individuals. Species dynamics are linked via predation mortality (M2) which varies with predator abundance, and size and species preference. Size preference is described with a preference function based upon a log-normal distribution and species preference with a diet matrix indicating who eats whom (Rochet et al. 2011, Thorpe et al. 2015). In each length class, individuals are also susceptible to residual natural mortality (M1) and fishing mortality (F). An ensemble approach is used, based upon a "filtered ensemble" (FE) of models drawn from a population of 78,125 candidate models (the "unfiltered ensemble" or UE), with the FE being selected on the basis of an individual member's ability to persist stocks when unfished, and to simulate assessed abundances of 10 stocks between 1990 and 2010 to an acceptable degree. This ensemble approach is described in detail in Thorpe et al. (2015), and further details of the model, including equations are

provided in Thorpe et al. (2017). The species in model are: sprat, Norway pout, sandeel, poor cod, long rough dab, dab, herring, horse mackerel, lemon sole, sole, mackerel, whiting, witch, gurnard, plaice, starry ray, haddock, cuckoo ray, monkfish, cod and saithe.

## Supplement 5 – Ensemble model and prior elicitation

### Model description

The predicted indicator values from the models and observations were combined using the ensemble model of Spence et al. (2018). Here we give a brief overview of the model.

At time t, the true indicator values, $\mathbf{y}^{(t)} = \left( y_1^{(t)}, \ldots, y_7^{(t)} \right)'$, was described by five models, $\widehat{\mathbf{x}_k^{(t)}} = \left( \widehat{x_{k,1}^{(t)}}, \ldots, \widehat{x_{k,n_k}^{(t)}} \right)'$, with $n_k$ outputs each relating to the yield of one of the species, for $k = 1, \ldots, 5$. In the ensemble model, uncertainty is separated into parameter, structural, and observation uncertainty.

Not all of the models output all of the indicators over the whole time period. For example, FishSums does not include zooplankton biomass and mizer spans the time-period 1984-2050 while EwE only covers the time period 1991-2050. To accommodate these differences, Spence et al. (2018) introduced a latent variable, known as the 'best guess', $\mathbf{x}_k^{(t)} = \left( x_1^{(t)}, \ldots, x_9^{(t)} \right)'$, which represents model $k$'s output if it were fitted to an infinite amount of data and was run with all of the indicators at time $t$. In this study, indicators are either present or absent in each model, therefore if the $k$th model was evaluated at time $t$, its output was

$$\widehat{\mathbf{x}_k^{(t)}} \sim N\left( M_k \mathbf{x}_k^{(t)}, \Sigma_k \right) \text{ (S8)}$$

where $M_k$ is $n_k \times 7$ matrix and $\Sigma_k$ reflects the parameter uncertainty of the $k$th model.

Structural uncertainty was considered by saying that the true value of the indicators at time $t$ was model $k$'s best guess plus a discrepancy term, $\zeta_k^{(t)}$ (Kennedy & O'Hagan 2001), i.e.

$$\mathbf{y}^{(t)} = \mathbf{x}_k^{(t)} + \boldsymbol{\zeta}_k^{(t)} \text{ (S9)}$$

The discrepancy term, $\boldsymbol{\zeta}_k^{(t)}$, is split between discrepancies that were shared between all of the models, and discrepancies that were specific to the $k$th model. These two discrepancies were further split into fixed discrepancies, the long-term shared discrepancy, $\boldsymbol{\delta}$, and model $k$'s long-term individual discrepancy, $\boldsymbol{\gamma}_k$, and dynamic discrepancies, the short-term shared discrepancy, $\boldsymbol{\eta}^{(t)}$, and model $k$'s short-term individual discrepancy, $\mathbf{z}_k^{(t)}$, i.e.

$$\boldsymbol{\zeta}_k^{(t)} = \boldsymbol{\delta} + \boldsymbol{\eta}^{(t)} + \boldsymbol{\gamma}_k + \mathbf{z}_k^{(t)} \text{ (S10)}$$

The long-term individual discrepancy for the $k$th model, $\boldsymbol{\gamma}_k$, was a random variable with an expectation of $0$ and covariance of $C$. The short-term discrepancy terms, $\boldsymbol{\eta}^{(t)}$ and $\mathbf{z}_k^{(t)}$ followed an auto-regressive processes of order one,

$$\boldsymbol{\eta}^{(t)} \sim N\left( R_\eta \boldsymbol{\eta}^{(t-1)}, \Lambda_\eta \right) \text{ (S11)}$$

and

$$\mathbf{z}_k^{(t)} \sim N\left( R_k \mathbf{z}_k^{(t-1)}, \Lambda_k \right) \text{ (S12)}$$

The dynamics of each model's short-term discrepancy were sampled from a hierarchical model with the $j$th diagonal elements of $R_i$ being

$$\frac{R_{i,j}+1}{2} \sim \text{Beta}(a_j, b_j) \quad \text{(S13)}$$

for the $i$th model and the off-diagonal elements being 0. The variance component from the short-term discrepancy of the $i$th model was parameterised by

$$\Lambda_i = \Pi_i P_i \Pi_i \quad \text{(S14)}$$

$\Pi_i$ is 7x7 matrix with the off-diagonal elements are zero and the $j$th diagonal element being

$$\pi_{i,j,j}^2 \sim \text{Gamma}(a_{\pi j}, b_{\pi j}) \quad \text{(S15)}$$

We specified permit separated distributions on for each element of $P_i$, subject to it being positive definite. The $j, k$th element of $P_i$ is

$$\frac{\rho_{ijk}+1}{2} \sim \begin{cases} \text{Beta}(a_{\rho jk}, b_{\rho jk}) & \text{if } j \neq k \\ 1 & \text{otherwise} \end{cases} \quad \text{(S16)}$$

Observation uncertainty captured noisy observations of the indicators, $\widehat{y^{(t)}}$, with

$$\widehat{y}^{(t)} \sim N(y^{(t)}, \Sigma_y) \quad \text{(S17)}$$

In this study $\Sigma_y$ was calculated from the polynomial regression that was used to fit the observations. The model specific values are described in Table S3 and a summary of the ensemble model is provided in Table S4. The elements of

$$y^{(t)} = (\ln ZB_t, \ln Z{:}P_t, \ln MML_t, \ln TyL_t, \text{logit } LFI_t, \ln BM_t, \ln TBF_t)' \quad \text{(S18)}$$

where the subscript $t$ means that it Is the true value of the indicator at time $t$.

Table S3. A summary of the models, their outputs, $M_i$, and a reference to the parameter uncertainty.

| Model | Description | Time period | Outputs | $M_i$ | Reference for $\Sigma_i$ |
|---|---|---|---|---|---|
| Ecopath with Ecosim | An ecosystem model with 60 functional groups for the North Sea | 1991-2050 | 1) Zooplankton biomass 2) Z:P ratio 3) BM biomass 4) Total fish biomass | $M_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ | Mackinson et al. (2018) |
| Mizer | A weight-based multi-species model for the North Sea | 1986-2050 | 1) MML 2) TyL 3) LFI 4) Total fish biomass | $M_2 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ | Spence et al. (2016) |
| FishSUMs | A length-based multi-species model for the North Sea | 1986-2050 | 1) MML 2) TyL 3) LFI 4) Total fish biomass | $M_3 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ | Spence et al. (2018) |

| | | | | | | |
|---|---|---|---|---|---|---|
| StrathE2E | An end-to-end ecosystem model for the North Sea | 1986-2050 | 1) Zooplankton biomass 2) Z:P ratio 3) BM biomass 4) Total fish biomass | $M_4 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ | | Spence et al. (2018) |
| LeMans | A length-based multi-species model for the North Sea | 1986-2050 | 1) MML 2) TyL 3) LFI 4) Total fish biomass | $M_5 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ | | Thorpe et al. (2015) |

Table S4. A summary of the ensemble model.

| Variable | Dimension | Time period | Description | Relationship |
|---|---|---|---|---|
| $\boldsymbol{y}^{(t)}$ | 7 | 1986-2050 | The true values of the indicators | $\boldsymbol{y}^{(t)} \sim N(\boldsymbol{y}^{(t-1)}, \Lambda_y)$ |
| $\widehat{\boldsymbol{y}}^{(t)}$ | 7 | 1986-2013 | Observations of the truth | $\widehat{\boldsymbol{y}}^{(t)} \sim N(\boldsymbol{y}^{(t)}, \Sigma_y)$ |
| $\boldsymbol{\delta}$ | 7 | NA | Long-term shared discrepancy | |
| $\boldsymbol{\eta}^{(t)}$ | 7 | 1986-2050 | Short-term shared discrepancy | $\boldsymbol{\eta}^{(t)} \sim N(R_\eta \boldsymbol{\eta}^{(t-1)}, \Lambda_\eta)$ |
| $\boldsymbol{\mu}^{(t)}$ | 7 | 1986-2050 | Model consensus | $\boldsymbol{\mu}^{(t)} = \boldsymbol{y}^{(t)} + \boldsymbol{\delta} + \boldsymbol{\eta}^{(t)}$ |
| $\boldsymbol{\gamma}_i$ | 7 | NA | Model $i$'s long-term individual discrepancy | $\boldsymbol{\gamma}_i \sim N(\boldsymbol{0}, C_\gamma)$ |
| $\boldsymbol{z}_i^{(t)}$ | 7 | 1986-2050 | Model $i$'s short-term individual discrepancy | $\boldsymbol{z}_i^{(t)} \sim N\left(R_i \boldsymbol{z}_i^{(t-1)}, \Lambda_i\right)$ |
| $\boldsymbol{x}_i^{(t)}$ | 7 | 1986-2050 | Model $i$'s best guess | $\boldsymbol{x}_i^{(t)} = \boldsymbol{\mu}^{(t)} + \boldsymbol{x}_i^{(t)} + \boldsymbol{z}_i^{(t)}$ |
| $\widehat{\boldsymbol{x}}_i^{(t)}$ | Model specific (see Table S3) | Model specific (see Table S3) | Model output | $\widehat{\boldsymbol{x}}_i^{(t)} \sim N\left(M_i \boldsymbol{x}_i^{(t)}, \Sigma_i\right)$ |

## Elicitation of prior distributions

The covariance of the long-term discrepancy was separated into the diagonal matrix $\Sigma_\gamma$, with values $\sigma_{\gamma i}$, giving the marginal standard deviations and the correlation matrix $P_\gamma$, with

$$C_\gamma = \Sigma_\gamma P_\gamma \Sigma_\gamma.$$

By applying the prior

$$p(P) = \mathbb{I}_{P>0} \prod_{i=1}^{n-1} \prod_{j=i+1}^{n} Beta(\rho_{ij}|a_{ij}, b_{ij}),$$

we are able to elicit experts' beliefs on correlations. Here $\mathbb{I}_{P>0}$ is an indicator function that takes the value 1 if P is positive definite and 0 otherwise, $\rho_{ij}$ is the element of $P$ on the $i$th row and $j$th column and $Beta(x|a, b)$ is the density of a $Beta(a, b)$ distribution evaluated at $x$.

We focus on the beliefs of one of the authors, MAS. Using the Shelf R package (Oakley 2020), MAS' prior beliefs were turned into probability distributions. The diagonal elements of $\Sigma_\gamma$, $\sigma_{\gamma i}$, were

$$\sigma_{\gamma i} \sim Exponential(5).$$

Using the method of concordance (Gokhale & James-Press 1982, Clemen & Reilly 1999), beta distributions were fitted to MAS' prior beliefs about the elements of the correlation matrix, $P$. The mean prior predictive distribution for $C_\gamma$ was

0.186 0.112 0.127 0.184 0.137 0.138 0.117

0.112 0.157 0.131 0.171 0.129 0.13 0.111

0.127 0.131 0.212 0.204 0.149 0.156 0.131

0.184 0.171 0.204 0.436 0.227 0.219 0.187

0.137 0.129 0.149 0.227 0.311 0.196 0.172

0.138 0.13 0.156 0.219 0.196 0.312 0.165

0.117 0.111 0.131 0.187 0.172 0.165 0.232.

The parameter values of $\Pi_i$ and $P_i$ were chosen in a similar way. The prior mean of the $\pi_{i,j}^2$ was 0.015 with a standard deviation of 0.030, for all $i, j$. The off-diagonal elements of $\Lambda_i$ had a mean of 0 and a standard deviation 0.015.

## Supplement 6 – Additional figures
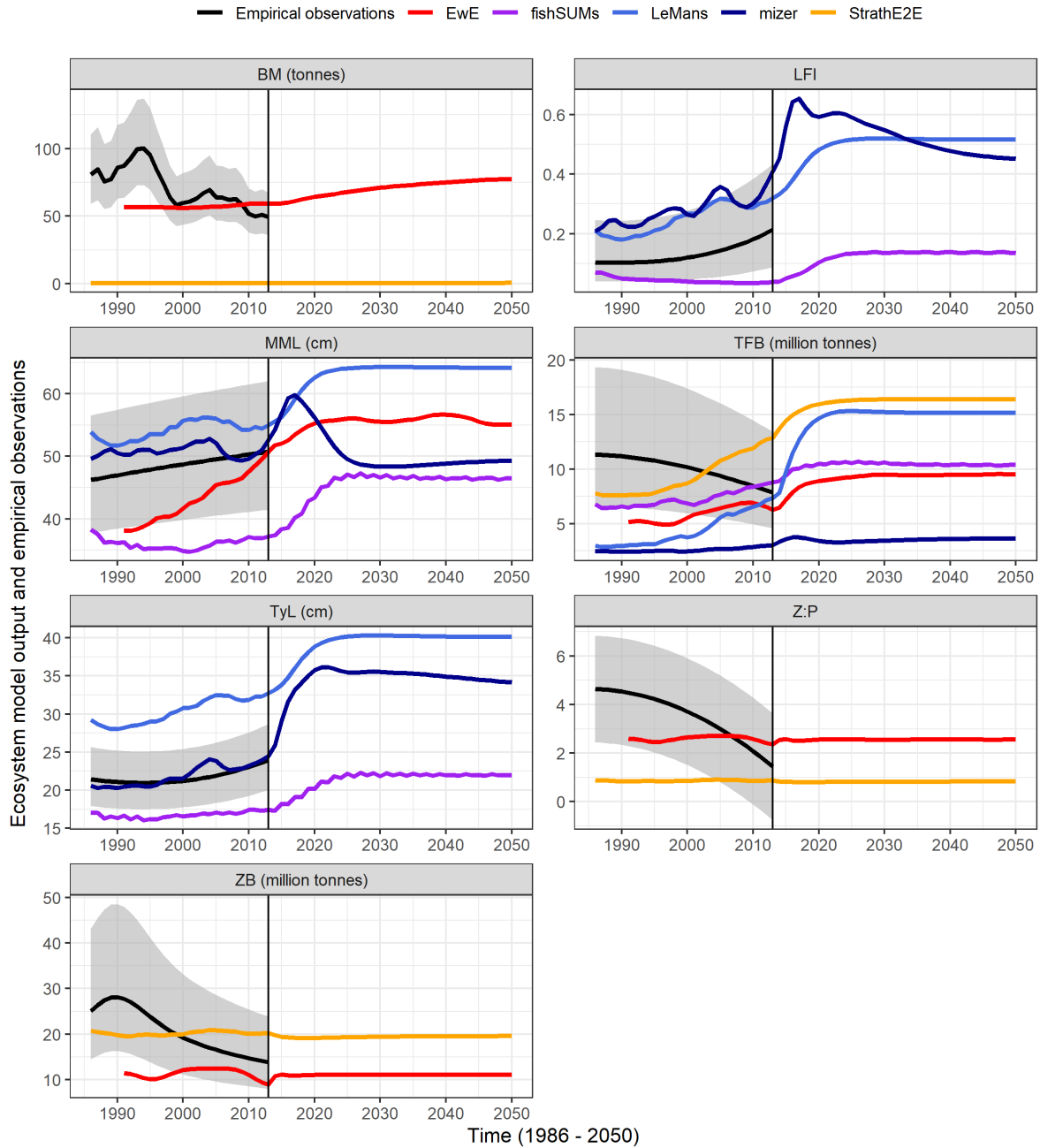


Figure S6. Indicator estimates from the combination of all five ecosystem models and observations until 2013. Starting in 2014, each ecosystem model is run under a no fishing scenario. The uncertainty envelope represents ± 2 standard deviations. The black vertical line indicates the last year for which observations were considered.
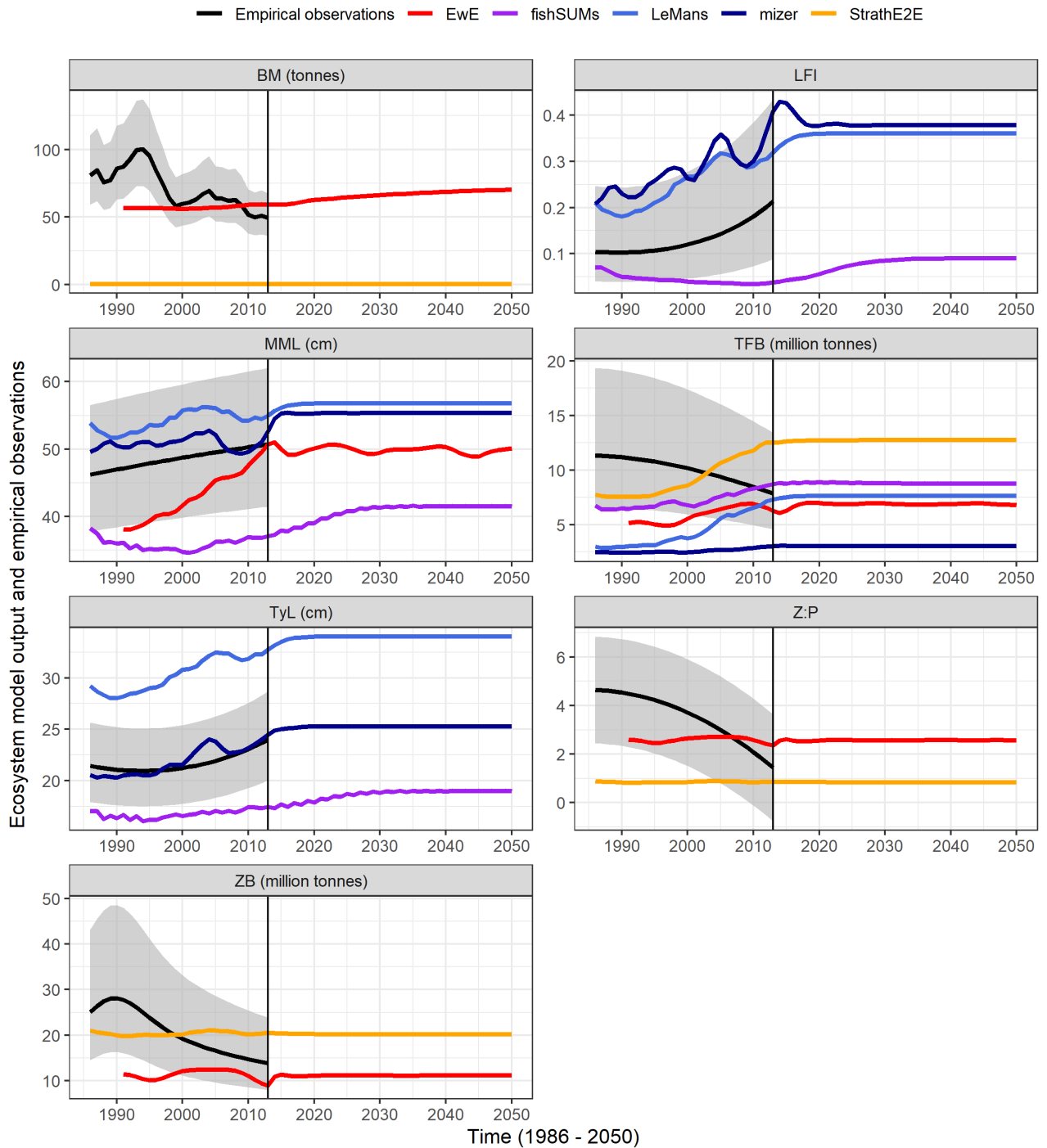
Figure S7. Indicator estimates from the combination of all five ecosystem models and observations until 2013. Starting in 2014, each ecosystem model is run under a status quo scenario. The uncertainty envelope represents ± 2 standard deviations. The black vertical line indicates the last year for which observations were considered.
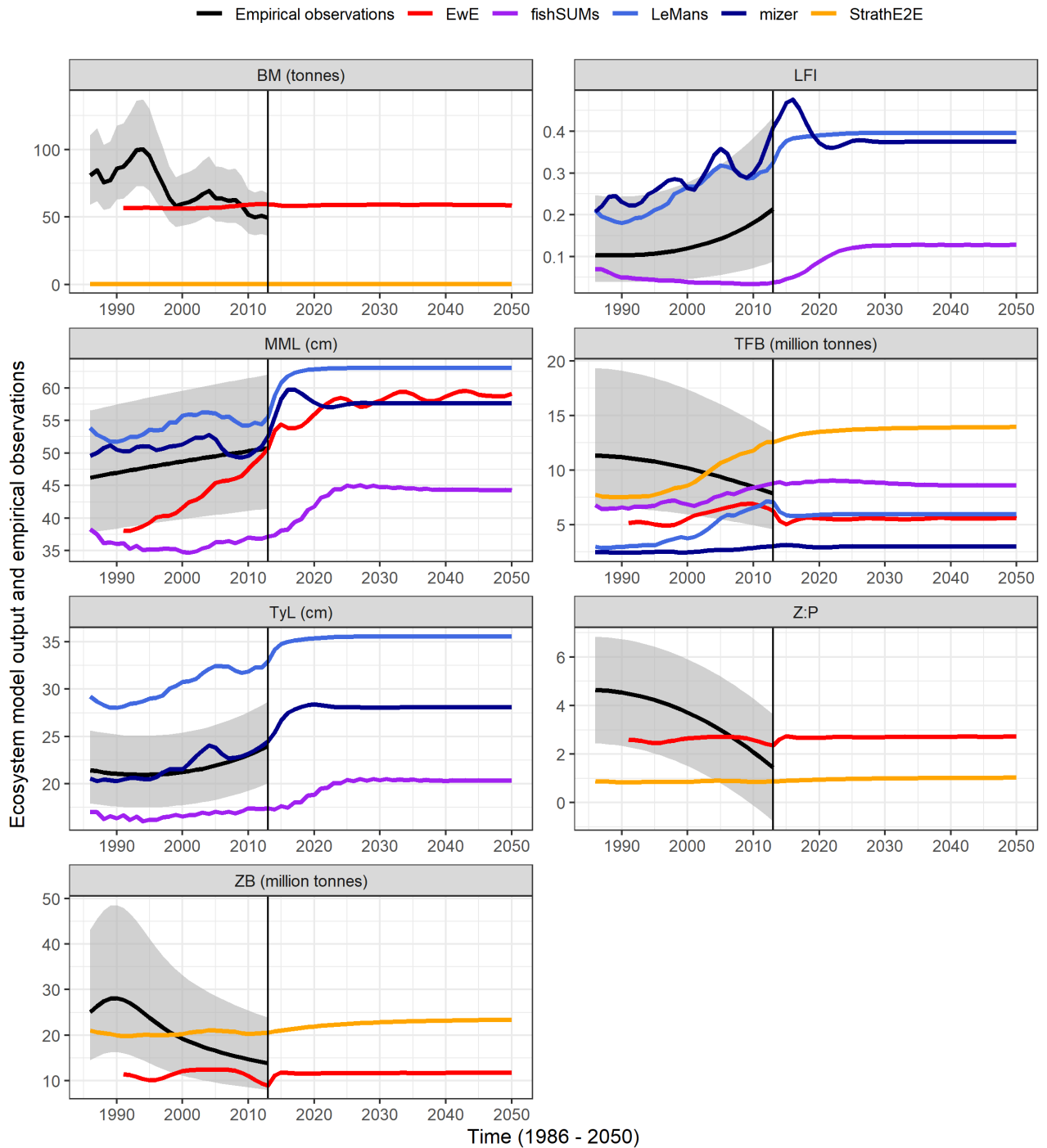
Figure S8. Indicator estimates from the combination of all five ecosystem models and observations until 2013. Starting in 2014, each ecosystem model is run under a NASH fishing scenario. The uncertainty envelope represents ± 2 standard deviations. The black vertical line indicates the last year for which observations were considered.
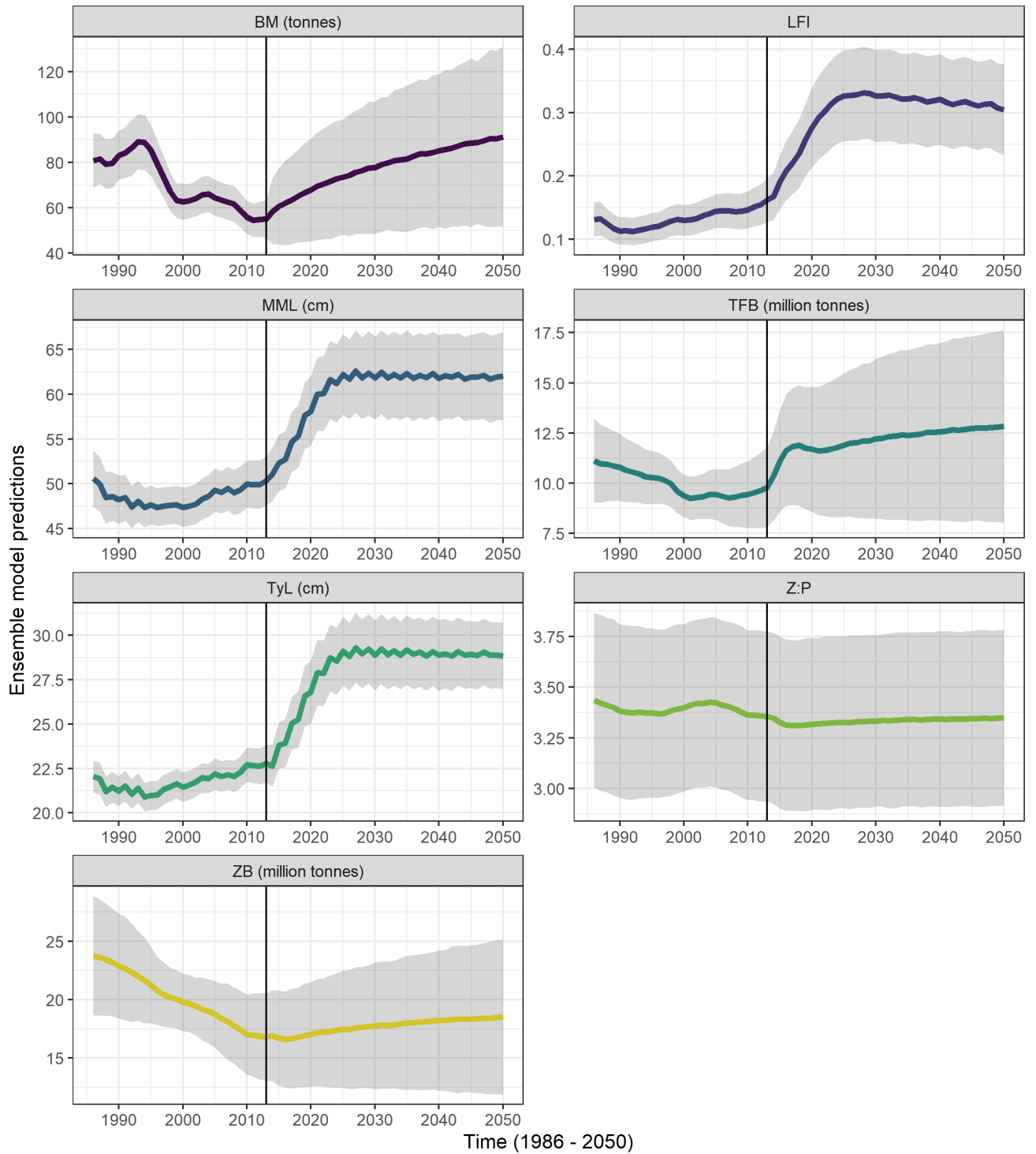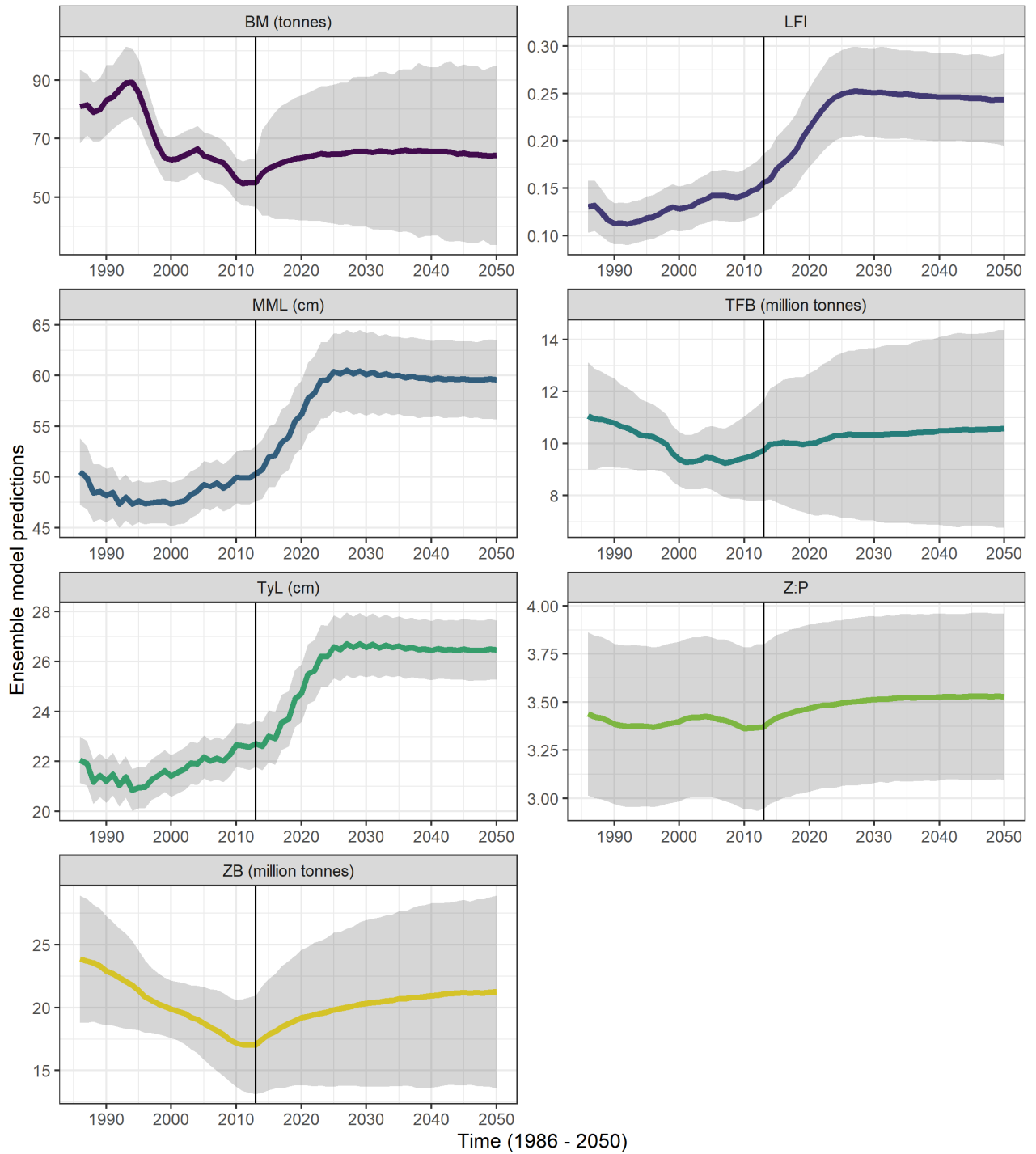
Figure S9. Indicator predictions from the ensemble model under a no fishing scenario. The uncertainty envelope surrounding each prediction represents ± 2 standard deviations. Note, the empirical observations end in 2013 as illustrated by the black vertical line.
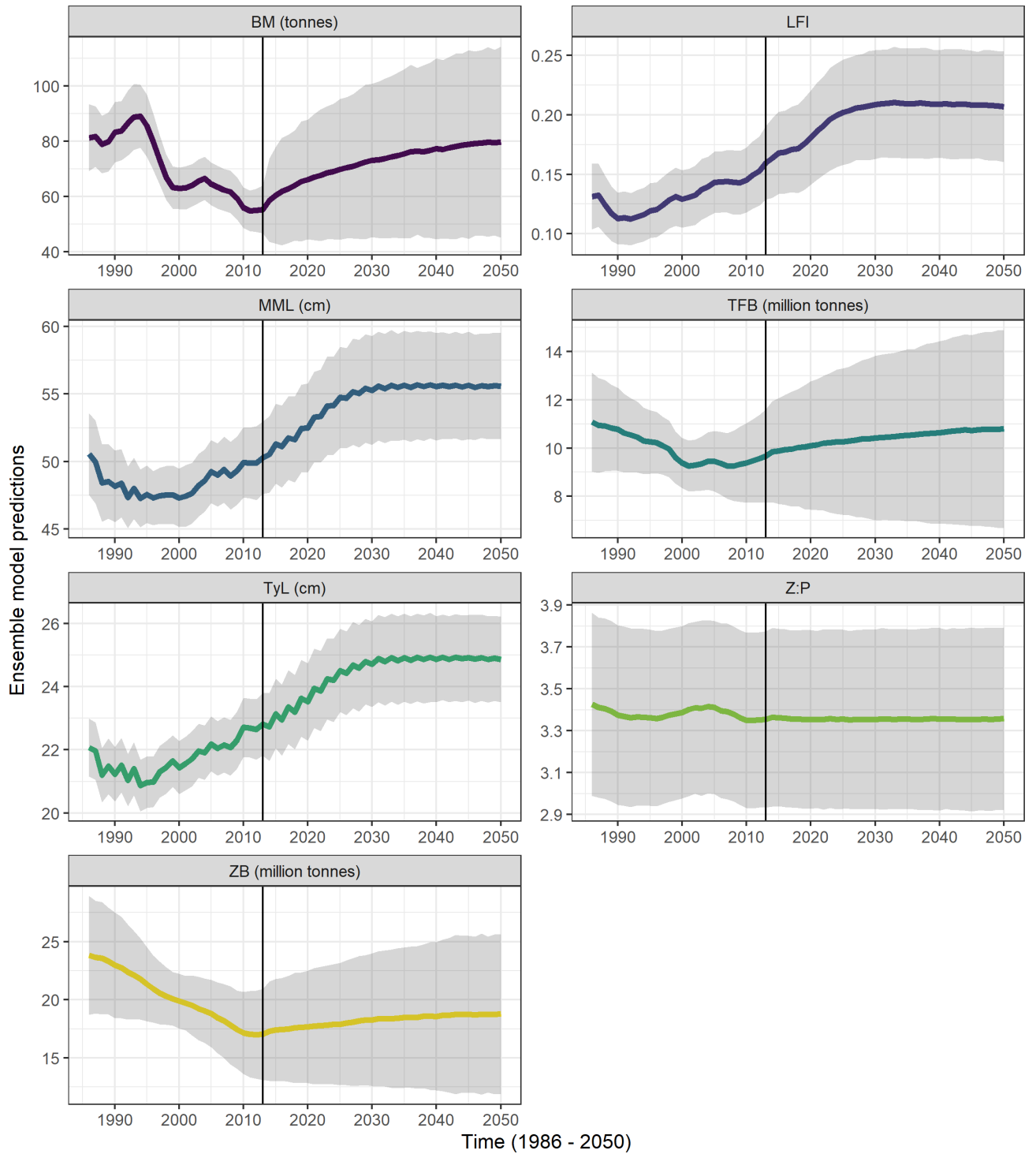
Figure S10. Indicator predictions from the ensemble model under a status quo scenario. The uncertainty envelope surrounding each prediction represents ± 2 standard deviations. Note, the empirical observations end in 2013 as illustrated by the black vertical line.

Figure S11. Indicator predictions from the ensemble model under a NASH fishing scenario. The uncertainty envelope surrounding each prediction represents ± 2 standard deviations. Note, the empirical observations end in 2013 as illustrated by the black vertical line.
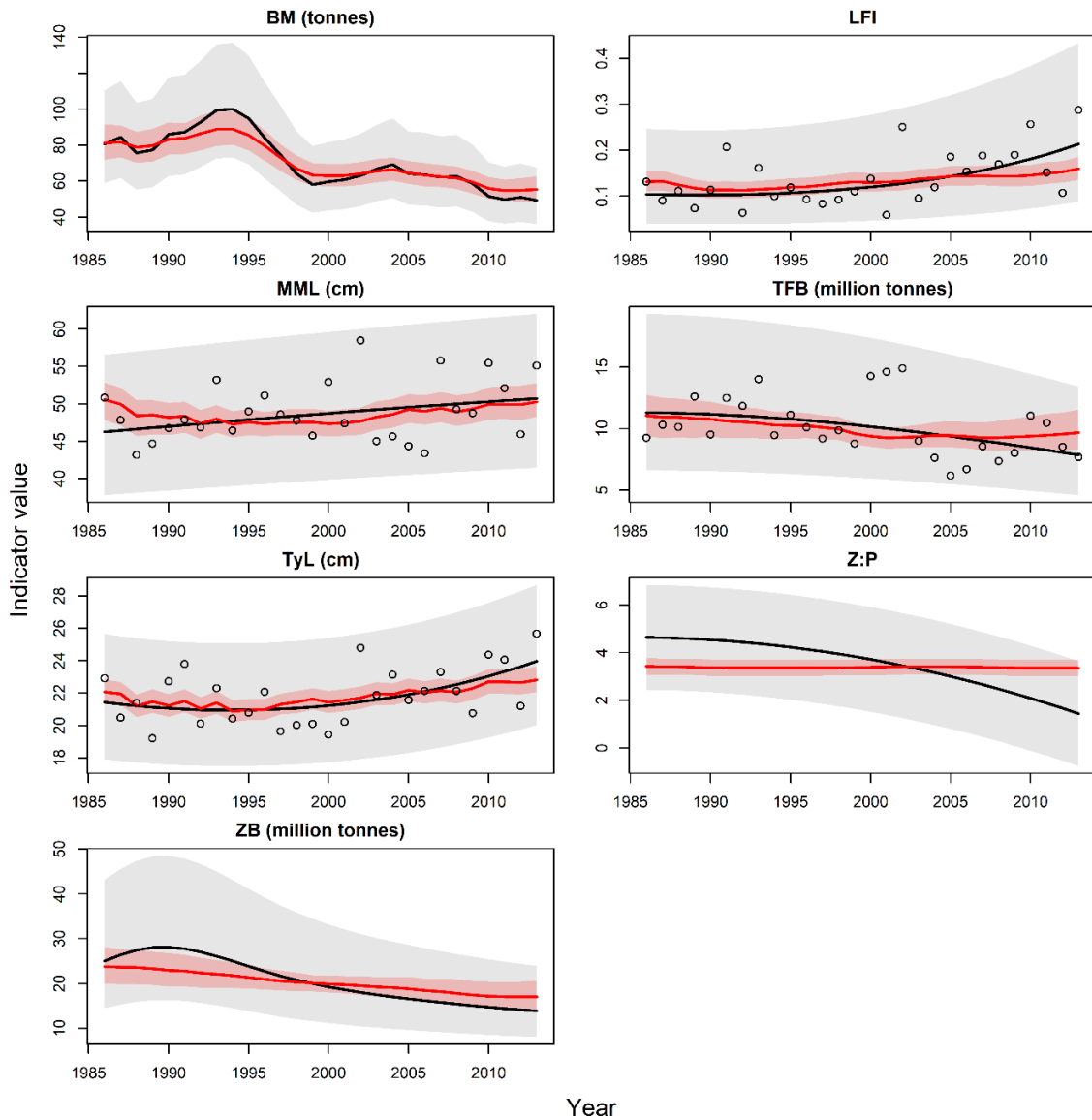
Figure S12. The median estimate of the empirical studies are shown by the black lines, with the 90% confidence intervals the shaded grey region, and the red line is the median estimate from the ensemble model, with the shaded red region being the 90% confidence interval. The points in the LFI, MML, TFB and TyL are the points from the IBTS survey (see Supplement 2).

## Literature Cited

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 doi:10.1111/1365-2664.12238

Christensen V, Walters CJ (2004) Ecopath with Ecosim: Methods, capabilities and limitations. Ecol Modell 172:109–139 doi:10.1016/j.ecolmodel.2003.09.003

Clemen RT, Reilly T (1999) Correlations and copulas for decision and risk analysis. Manage Sci 45:208–224 doi:10.1287/mnsc.45.2.208

Doucet A, Johansen A (2008) A Tutorial on Particle Filtering and Smoothing. -unknown>

Elith J, Leathwick JR (2009) Species distribution models: Ecological explanation and prediction across space and time. Annu Rev Ecol Evol Syst 40:677–697 doi:10.1146/annurev.ecolsys.110308.120159

Evans PGH, Waggitt JJ (2020) The impacts of climate change on fisheries, relevant to the coastal and marine environment around the UK. Mar Clim Chang Impacts Partnersh Sci Rev:421–455.

Gokhale DV, James-Press S (1982) Assessment of a Prior Distribution for the Correlation Coefficient in a Bivariate Normal Distribution. J R Stat Soc [Ser A] 145:237–249 doi:10.2307/2981537

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 doi:10.1139/f06-039

Hartvig M, Andersen KH, Beyer JE (2011) Food web framework for size-structured populations. J Theor Biol 272:113–122 PubMed doi:10.1016/j.jtbi.2010.12.006

Heath MR (2012) Ecosystem limits to food web fluxes and fisheries yields in the North Sea simulated with an end-to-end food web model. Prog Oceanogr 102:42–66 doi:10.1016/j.pocean.2012.03.004

Heath MR, Cook RM, Cameron AI, Morris DJ, Speirs DC (2014a) Cascading ecological effects of eliminating fishery discards. Nat Commun 13:3893 PubMed

Heath MR, Speirs DC, Steele JH (2014b) Understanding patterns and processes in models of trophic cascades. Ecol Lett 17:101–114 PubMed doi:10.1111/ele.12200

Heath MR, Wilson R, Speirs DC (2015) Modelling the Whole-Ecosystem Impacts of Trawling. Glasgow, UK.

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 doi:10.1016/j.ecolmodel.2015.12.007

John EH, Batten SD, Stevens D, Walne AW, Jonas T, Hays GC (2002) Continuous plankton records stand the test of time: Evaluation of flow rates, clogging and the continuity of the CPR time-series. J Plankton Res 24:941–946 doi:10.1093/plankt/24.9.941

Kennedy MC, O'Hagan A (2001) Bayesian calibration of computer models. J R Stat Soc Series B Stat Methodol 63:425–464 doi:10.1111/1467-9868.00294

Lynam CP, Mackinson S (2015) How will fisheries management measures contribute towards the attainment of Good Environmental Status for the North Sea ecosystem? Glob Ecol Conserv 4:160–175 doi:10.1016/j.gecco.2015.06.005

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):e0190015 PubMed doi:10.1371/journal.pone.0190015

Mitchell I, Daunt F, Frederiksen M, Wade K (2020) Impacts of climate change on seabirds, relevant to the coastal and marine environment around the UK. Mar Clim Chang Impacts Partnersh Sci Rev:382–399.

Morris DJ, Speirs DC, Cameron AI, Heath MR (2014) Global sensitivity analysis of an end-to-end marine ecosystem model of the North Sea: Factors affecting the biomass of fish and benthos. Ecol Modell 273:251–263 doi:10.1016/j.ecolmodel.2013.11.019

Oakley J (2020) Tools to Support the Sheffield Elicitation Framework (SHELF R Package). 1–44.

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 doi:10.1016/j.icesjms.2006.03.009

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 doi:10.1007/BF00306135

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 doi:10.1139/F10-159

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 PubMed doi:10.1111/2041-210X.12256

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 doi:10.1016/j.fishres.2010.09.023

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 doi:10.1016/j.ecolmodel.2015.10.032

Spence MA, Blackwell PG, Blanchard JL (2016) Parameter uncertainty of a dynamic multispecies size spectrum model. Can J Fish Aquat Sci 73:589–597 doi:10.1139/cjfas-2015-0022

Spence MA, Blanchard JL, Rossberg AG, Heath MR and others (2018) A general framework for combining ecosystem models. Fish Fish 19:1031–1042 doi:10.1111/faf.12310

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 PubMed doi:10.1111/2041-210X.12292

Thorpe RB, Dolder PJ, Reeves S, Robinson P, Jennings S (2016) Assessing fishery and ecological consequences of alternate management options for multispecies fisherie. ICES J Mar Sci Sci 73:1503–1512 doi:10.1093/icesjms/fsw028

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 doi:10.1093/icesjms/fsx062

Waggitt JJ, Evans PGH, Andrade J, Banks AN and others (2020) Distribution maps of cetacean and seabird populations in the North-East Atlantic. J Appl Ecol 57:253–269 doi:10.1111/1365-2664.13525