

# Multispecies consequences of fishing in the North Sea

Michael A. Spence<sup>1\*</sup> and Matthew R. Kerr<sup>1</sup>

<sup>1</sup>Centre for Environment, Fisheries and Aquaculture Science, Pakefield Road,  
Lowestoft, Suffolk NR33 0HT, UK

\*michael.spence@cefas.co.uk

## 1 Introduction

This document outlines the methods used to calculate the long-term spawning stock biomass (SSB),  $B_i(\mathbf{F})$ , of nine species in the North Sea (see Table 1) under fixed fishing mortality from 2020,  $\mathbf{F} = (F_1, \dots, F_9)'$ , with  $F_i \in [0, 2]$  for  $i = 1, \dots, 9$ .

Table 1: A summary of the species in the model.

$i$	Species	Latin name
1	Sandeel	<i>Ammodytes spp.</i>
2	Norway pout	<i>Trisopterus esmarkii</i>
3	Herring	<i>Clupea harengus</i>
4	Whiting	<i>Merlangius merlangus</i>
5	Sole	<i>Solea solea</i>
6	Plaice	<i>Pleuronectes platessa</i>
7	Haddock	<i>Melanogrammus aeglefinus</i>
8	Cod	<i>Gadus morhua</i>
9	Saithe	<i>Pollachius virens</i>

To predict the long-term biomass, with quantifiable uncertainty, we used four multispecies models, henceforth known as simulators (to avoid confusion with the ensemble model), and stock-assessments, combined together using the ensemble model of Spence *et al.* (2018). The ensemble model was able to calculate the probability by breaking uncertainty down into parameter uncertainty, structural uncertainty, observational uncertainty and functional uncertainty. Functional uncertainty is the uncertainty in the simulators predictions for a fishing mortality that they had not been evaluated at and is a consequence of not being able to evaluate the simulators at all fishing mortalities.

The rest of this document is as follow: Section 2 describes the simulators used in the study, Section 3 sets out the ensemble model. We use a Gaussian process emulator to describe the functional uncertainty in Section 4, with the distribution of the long-term SSB and risk of falling below reference points being described in Section 5.

## 2 Simulators

We used four simulators to predict SSB: EcoPath with EcoSim (EwE Mackinson *et al.*, 2018), LeMans (Thorpe *et al.*, 2015), mizer (Blanchard *et al.*, 2014) and FishSUMs (Speirs *et al.*, 2016). All of the simulators were able to describe the dynamics of all nine species with the exception of FishSUMs, which did not include sole.

Fishing mortality was used to drive the dynamics of the simulators in the past (1984–2019). Typically fishing mortality is calibrated to data and therefore the interpretation of fishing mortality will be specific to the model used to calibrate them (Spence *et al.*, 2021; Skogen *et al.*, 2021). To keep the interpretation of fishing mortality across the study, we used the single-species assessments' fishing mortality at age to drive the dynamics of each of the simulators (ICES, 2020a,b). In the future (2020–2100), we assume that the age selectivity will be the same as in 2019 and species that appear in the models but not in the study will be fished at their 2019 levels. For the length and size-based simulators, LeMans, mizer and FishSUMs, we calculated the size at age using their respective von Bertalanffy parameters, however for EwE we used the  $\bar{F}$  values from the assessments. Each of the simulators was run using historical fishing mortality (prior to 2020) and then projected with fixed fishing mortality  $F$ , to 2100. Each of the simulators were developed and calibrated in a previous study. Table 2 contains a brief description of each of the models and a reference to their formulation and calibration.

Table 2: A summary of the simulators, including a reference to the model, their outputs used in the case study, the simulator-specific values of  $n_k$ ,  $T_k$ ,  $M_k$  and  $\Sigma_k$ .  $I_9$  is a 9 dimensional indicator matrix.

$k$	Simulator	Description	$n_k$	$T_k$	$M_k$	Reference for $\Sigma_k$
1	EcoPath with EcoSim (EwE)	An ecosystem model with 60 functional groups for the North Sea (ICES, 2016).	$n_1 = 9$	$T_1 = 1991 - 2100$	$M_1 = I_9$	Mackinson <i>et al.</i> (2018)
2	LeMans	A discrete time length-based model that describes growth and predation (Thorpe <i>et al.</i> , 2015).	$n_2 = 9$	$T_2 = 1986 - 2100$	$M_2 = I_9$ .	Thorpe <i>et al.</i> (2015)
3	mizer	A size-based model that describes ontogenetic feeding and growth, mortality, and reproduction driven by size-dependent predation and maturation processes (Blanchard <i>et al.</i> , 2014).	$n_3 = 9$	$T_3 = 1984 - 2100$	$M_3 = I_9$	Spence <i>et al.</i> (2016)

4 FishSUMs

A discrete time length-based model that describes growth, density-dependent mortality, and losses due to fishing and predation by explicitly modelled species, and seasonal reproduction (Speirs *et al.*, 2016).

$n_4 = 8$        $T_4 = 1984 - 2100$ .

Spence *et al.* (2018)

$$M_4 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

### 3 Ensemble model

At time  $t$ , the true SSB of the nine species under fishing scenario  $\mathbf{F}$ ,  $\mathbf{y}^{(t)}(\mathbf{F}) = (y_1^{(t)}(\mathbf{F}), \dots, y_9^{(t)}(\mathbf{F}))'$ , is described by four simulators,  $\hat{\mathbf{x}}_k^{(t)}(\mathbf{F}) = (\hat{x}_{k,1}^{(t)}(\mathbf{F}), \dots, \hat{x}_{k,n_k}^{(t)}(\mathbf{F}))'$ , with  $n_k$  outputs each relating to the SSB of one of the species, for  $k = 1, \dots, 4$ , and by assessments,  $\hat{\mathbf{y}}^{(t)} = (\hat{y}_1^{(t)}, \dots, \hat{y}_9^{(t)})'$ .

Not all of the simulators output all nine species over the whole time period. For example, FishSums does not include common sole, 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’,  $\mathbf{x}_k^{(t)}(\mathbf{F}) = (x_1^{(t)}(\mathbf{F}), \dots, x_9^{(t)}(\mathbf{F}))'$ , which represents simulator  $k$ ’s output if it described all nine species at time  $t$  with no parameter uncertainty. In this study, species are either present or absent in each simulator, therefore if the  $k$ th simulator was evaluated at time  $t$ , its output was

$$\hat{\mathbf{x}}_k^{(t)}(\mathbf{F}) \sim N(M_k \mathbf{x}_k^{(t)}(\mathbf{F}), \Sigma_k), \quad (1)$$

where  $M_k$  is a  $n_k \times 9$  matrix and  $\Sigma_k$  reflects the parameter uncertainty of the  $k$ th simulator. A reference to where it was calculated for each simulator can be found in Table 2.

Structural uncertainty was considered by saying that the true value of the yield at time  $t$  is simulator  $k$ ’s best guess plus a discrepancy term,  $\zeta_k^{(t)}(\mathbf{F})$  (Kennedy and O’Hagan, 2001), i.e.

$$\mathbf{y}^{(t)}(\mathbf{F}) = \mathbf{x}_k^{(t)}(\mathbf{F}) + \zeta_k^{(t)}(\mathbf{F}). \quad (2)$$

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

$$\zeta_k^{(t)}(\mathbf{F}) = \boldsymbol{\delta} + \boldsymbol{\eta}^{(t)}(\mathbf{F}) + \boldsymbol{\gamma}_k + \mathbf{z}_k^{(t)}(\mathbf{F}). \quad (3)$$

The long-term individual discrepancy for the  $k$ th simulator is

$$\boldsymbol{\gamma}_k \sim N(\mathbf{0}, C_\gamma). \quad (4)$$

The short-term discrepancy terms,  $\boldsymbol{\eta}^{(t)}(\mathbf{F})$  and  $\mathbf{z}_k^{(t)}(\mathbf{F})$ , follow an auto-regressive processes of order one,

$$\boldsymbol{\eta}^{(t)}(\mathbf{F}) \sim N(R_\eta \boldsymbol{\eta}^{(t-1)}(\mathbf{F}), \Lambda_\eta) \quad (5)$$

and

$$\mathbf{z}_k^{(t)}(\mathbf{F}) \sim N(R_k \mathbf{z}_k^{(t-1)}(\mathbf{F}), \Lambda_k) \quad (6)$$

respectively.

In the absence of any other information, we expect the true biomass to evolve according to a random walk,

$$\mathbf{y}^{(t)}(\mathbf{F}) \sim N(\mathbf{y}^{(t-1)}(\mathbf{F}), \Lambda_y), \quad (7)$$

with, noisy observations of the SSB,

$$\hat{\mathbf{y}}^{(t)} \sim N(\mathbf{y}^{(t)}(\mathbf{F}), \Sigma_y). \quad (8)$$

In this study, the diagonal elements of  $\Sigma_y$  were calculated from uncertainty in the assessment, and the off-diagonal elements to zero (ICES, 2020a,b).

A summary of the ensemble model can be found in Table 3 and the simulator specific values are described in Table 2. For more details on the model see Spence *et al.* (2018). Due to the high dimensionality and correlation of the uncertain parameter space, the ensemble model was fitted using the No U-turn Hamiltonian Monte Carlo (Hoffman and Gelman, 2011) in the package Stan (Stan Development Team, 2020). We ran the algorithm for 2000 iterations discarding the first 1000 as burn in.

Table 3: A summary of the variables in the ensemble model. The ensemble model is run for 1984–2050. For values of  $n_k$ ,  $M_k$  and  $T_k$  see Table 2.

Variable	Dimensions	$t$	Description	Relationship
$\mathbf{y}^{(t)}$	9	1984–2100	The true SSB	$\mathbf{y}^{(t)} \sim N(\mathbf{y}^{(t-1)}, \Lambda_y)$
$\hat{\mathbf{y}}^{(t)}$	9	1984–2019	Noisy observation of $\mathbf{y}^{(t)}$	$\hat{\mathbf{y}}^{(t)} \sim N(\mathbf{y}^{(t)}, \Sigma_y)$
$\delta$	9	NA	Long-term shared discrepancy	
$\boldsymbol{\eta}^{(t)}$	9	1984–2100	Short-term shared discrepancy	$\boldsymbol{\eta}^{(t)} \sim N(R_\eta \boldsymbol{\eta}^{(t-1)}, \Lambda_\eta)$
$\boldsymbol{\mu}^{(t)}$	9	1984–2100	Simulator consensus	$\boldsymbol{\mu}^{(t)} = \mathbf{y}^{(t)} + \delta + \boldsymbol{\eta}^{(t)}$
$\boldsymbol{\gamma}_k$	9	NA	Simulator $k$ 's long-term individual discrepancy	$\boldsymbol{\gamma}_k \sim N(\mathbf{0}, C_\gamma)$
$\mathbf{z}_k^{(t)}$	9	1984–2100	Simulator $k$ 's short-term individual discrepancy	$\mathbf{z}_k^{(t)} \sim N(R_k \mathbf{z}_k^{(t-1)}, \Lambda_k)$
$\mathbf{x}_k^{(t)}$	9	1984–2100	Simulator $k$ 's best guess	$\mathbf{x}_k^{(t)} = \boldsymbol{\mu}^{(t)} + \boldsymbol{\gamma}_k + \mathbf{z}_k^{(t)}$
$\hat{\mathbf{x}}_k^{(t)}$	$n_k$	$T_k$	The expectation of simulator $k$ 's output $\mathbf{x}_k^{(t)}$	$\hat{\mathbf{x}}_k^{(t)} \sim N(M_k \mathbf{x}_k^{(t)}, \Sigma_k)$

## 4 Emulators

Each of the simulators was run to approximately a steady state by 2100,

$$\hat{\mathbf{x}}_k(\mathbf{F}) = \lim_{t \rightarrow \infty} \hat{\mathbf{x}}_k^{(t)}(\mathbf{F}) \approx \hat{\mathbf{x}}_k^{(2100)}(\mathbf{F}). \quad (9)$$

The difference between the two was very small so we were able to assume the approximation was valid. The simulators ran  $m$  future fishing scenarios to find  $\hat{\mathbf{x}}_k(\mathbf{F}^{(l)})$  for  $l = 1, \dots, m$  and  $k = 1, \dots, 4$ . We used a Gaussian process (Kennedy and O'Hagan, 2001; Noè *et al.*, 2019) to emulate  $\hat{\mathbf{x}}_k(\mathbf{F})$  for all values of  $\mathbf{F}$ . We define  $\mathbf{f}_{k,i} = (\hat{x}_{k,i}(\mathbf{F}^{(1)}), \hat{x}_{k,i}(\mathbf{F}^{(2)}), \dots, \hat{x}_{k,i}(\mathbf{F}^{(m)}))'$ , the value of the  $i$ th species for all  $\mathbf{F}$  values that have been evaluated by the ensemble model and by the  $k$ th simulator for  $i = 1, \dots, 9$  and  $k = 1, \dots, 4$ , with

$$\mathbf{f}_{k,i} \sim GP(\boldsymbol{\psi}_{k,i}, K_{k,i}), \quad (10)$$

where  $\boldsymbol{\psi}_{k,i}$  was a generalised additive model (Wood, 2017) and  $K_{k,i}$  was the Matérn covariance function, fitted using the DiceKriging package (Roustant *et al.*, 2012) in R (R Core Team, 2020).

For a fishing scenario  $\mathbf{F} \notin \{\mathbf{F}^{(1)}, \dots, \mathbf{F}^{(m)}\}$  then

$$\hat{x}_{k,i}(\mathbf{F}) \sim N(\mu_{k,i}(\mathbf{F}), \xi_{k,i}(\mathbf{F})^2), \quad (11)$$

with  $\mu_{k,i}(\mathbf{F})$  and  $\xi_{k,i}(\mathbf{F})^2$  calculated from equation 10. If the  $k$ th simulator did not model the  $i$ th species then  $\xi_{k,i}(\mathbf{F})^2$  was made large so that the simulator effectively had zero precision.

## 5 Long-term SSB

The distribution of the long-term SSB, given the simulator runs and the stock assessment,  $E$ , was

$$p(\mathbf{B}(\mathbf{F})|E) = \int p(\mathbf{B}(\mathbf{F})|\boldsymbol{\theta})p(\boldsymbol{\theta}|E)d\boldsymbol{\theta}, \quad (12)$$

where  $\mathbf{B}(\mathbf{F}) = (B_1(\mathbf{F}), \dots, B_9(\mathbf{F}))'$  and

$$\boldsymbol{\theta} = \{\Lambda_y, \Lambda_\eta, R_\eta, \boldsymbol{\delta}, \boldsymbol{\gamma}_{1:4}, \Lambda_{1:4}, R_{1:4}\}. \quad (13)$$

This equation is not tractable, as it is not possible to calculate  $p(\boldsymbol{\theta}|E)$  exactly, however we used MCMC to sample 1000 times from this. Therefore,

$$p(\mathbf{B}(\mathbf{F})|E) \approx \frac{1}{1000} \sum_{j=1}^{1000} p(\mathbf{B}(\mathbf{F})|\boldsymbol{\theta}_j) \quad (14)$$

using Monte Carlo integration. For a single parameter set  $\boldsymbol{\theta}$ , then

$$\mathbf{B}(\mathbf{F})|\boldsymbol{\theta} = \lim_{t \rightarrow \infty} \mathbf{y}^{(t)}(\mathbf{F}) \sim N(\boldsymbol{\tau}(\mathbf{F}), S), \quad (15)$$

where

$$S = \Gamma_\eta + \left( \sum_{k=1}^4 D_k^{-1} \right)^{-1}, \quad (16)$$

with  $D_k = \Sigma_k + \Gamma_k + \Xi_k$ , with  $\Xi_k$  a  $9 \times 9$  matrix with  $\xi_{k,i}(\mathbf{F})^2$  on the  $i$ th diagonal and 0s everywhere else,

$$\text{vec}(\Gamma_\eta) = (\mathbb{I} - R_\eta \otimes R_\eta)^{-1} \text{vec}(\Lambda_\eta), \quad (17)$$

$$\text{vec}(\Gamma_i) = (\mathbb{I} - R_i \otimes R_i)^{-1} \text{vec}(\Lambda_i) \quad (18)$$

and

$$\boldsymbol{\tau}(\mathbf{F}) = \left( \sum_{k=1}^4 D_k^{-1} \right)^{-1} \left( \sum_{k=1}^4 D_k^{-1} (\boldsymbol{\mu}_k(\mathbf{F}) - \boldsymbol{\gamma}_k) \right) - \boldsymbol{\delta}, \quad (19)$$

with  $\boldsymbol{\mu}_k(\mathbf{F}) = (\mu_{k,1}(\mathbf{F}), \dots, \mu_{k,9}(\mathbf{F}))$ .

## 5.1 Risk

Risk for the  $i$ th species is defined as the probability that the long-term SSB is below calculated the reference point,  $r_i$ , i.e

$$Pr(B_i(\mathbf{F}) < r_i|E) = \int Pr(B_i(\mathbf{F}) < r_i|\boldsymbol{\theta})p(\boldsymbol{\theta}|E)d\boldsymbol{\theta}. \quad (20)$$

Like equation 12, equation 20 is not tractable and we use Monte Carlo integration to approximate it,

$$Pr(B_i(\mathbf{F}) < r_i|E) \approx \frac{1}{1000} \sum_{j=1}^{1000} Pr(B_i(\mathbf{F}) < r_i|\boldsymbol{\theta}_j), \quad (21)$$

with  $\boldsymbol{\theta}_j$  coming from the MCMC. As we know from equation 15

$$B_i(\mathbf{F})|\boldsymbol{\theta} \sim N(\tau_i(\mathbf{F}), S_{ii}), \quad (22)$$

therefore

$$Pr(B_i(\mathbf{F}) < r_i|\boldsymbol{\theta}) = \Phi\left(\frac{\tau_i(\mathbf{F}) - r_i}{\sqrt{S_{ii}}}\right), \quad (23)$$

where  $\Phi(\cdot)$  is the cumulative distribution of the standard normal distribution.

## References

- Blanchard, J. L., Andersen, K. H., Scott, F., Hintzen, N. T., Piet, G., and Jennings, S. (2014). Evaluating targets and trade-offs among fisheries and conservation objectives using a multispecies size spectrum model. *Journal of Applied Ecology*, **51**(3), 612–622.
- Hoffman, M. and Gelman, A. (2011). The no-u-turn sampler: Adaptively setting path lengths in hamiltonian monte carlo. *Journal of Machine Learning Research*, **15**.
- ICES (2016). Working Group on Multispecies Assessment Methods (WGSAM). Technical report, International Council for Exploration of the Seas.
- ICES (2020a). Herring Assessment Working Group for the Area South of 62 N (HAWG). Technical report, ICES Scientific Reports. 2:60. 1151 pp, ICES, Copenhagen.
- ICES (2020b). Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak. Technical report, ICES Scientific Reports. 2:61. 1353, ICES, Copenhagen.
- Kennedy, M. C. and O’Hagan, A. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **63**(3), 425–464.
- Mackinson, S., Platts, M., Garcia, C., and Lynam, C. (2018). Evaluating the fishery and ecological consequences of the proposed North Sea multi-annual plan. *PLOS ONE*, **13**(1), 1–23.
- Noè, U., Lazarus, A., Gao, H., Davies, V., Macdonald, B., Mangion, K., Berry, C., Luo, X., and Husmeier, D. (2019). Gaussian process emulation to accelerate parameter estimation in a mechanical model of the left ventricle: a critical step towards clinical end-user relevance. *Journal of The Royal Society Interface*, **16**(156), 20190114.



- R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Roustant, O., Ginsbourger, D., and Deville, Y. (2012). DiceKriging, DiceOptim: Two R packages for the analysis of computer experiments by kriging-based metamodeling and optimization. *Journal of Statistical Software*, **51**(1), 1–55.
- Skogen, M. D., Ji, R., Akimova, A., Daewel, U., Hansen, C., Hjøllø, S. S., van Leeuwen, S. M., Maar, M., Macias, D., Mousing, E. A., Almroth-Rosell, E., Saille, S. F., Spence, M. A., Troost, T. A., and van de Wolfshaar, K. (2021). Disclosing the truth: Are models better than observations? *Marine Ecology Progress Series*, **dynmod**, DYNMODav1.
- Speirs, D., Greenstreet, S., and Heath, M. (2016). Modelling the effects of fishing on the North Sea fish community size composition. *Ecological Modelling*, **321**, 35–45.
- Spence, M. A., Blackwell, P. G., and Blanchard, J. L. (2016). Parameter uncertainty of a dynamic multispecies size spectrum model. *Canadian Journal of Fisheries and Aquatic Sciences*, **73**(4), 589–597.
- Spence, M. A., Blanchard, J. L., Rossberg, A. G., Heath, M. R., Heymans, J. J., Mackinson, S., Serpetti, N., Speirs, D. C., Thorpe, R. B., and Blackwell, P. G. (2018). A general framework for combining ecosystem models. *Fish and Fisheries*, **19**(6), 1031–1042.
- Spence, M. A., Thorpe, R. B., Blackwell, P. G., Scott, F., Southwell, R., and Blanchard, J. L. (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 and Fisheries*, **22**(4), 667–681.
- Stan Development Team (2020). RStan: the R interface to Stan. R package version 2.19.3.
- Thorpe, R. B., Le Quesne, W. J. F., Luxford, F., Collie, J. S., and Jennings, S. (2015). Evaluation and management implications of uncertainty in a multispecies size-structured model of population and community responses to fishing. *Methods in Ecology and Evolution*, **6**(1), 49–58.
- Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC, Second Edition edition.