WORKING GROUP ON MULTISPECIES ASSESSMENT METHODS (WGSAM)

VOLUME 6 | ISSUE 99

ICES SCIENTIFIC REPORTS

RAPPORTS SCIENTIFIQUES DU CIEM

ICES INTERNATIONAL COUNCIL FOR THE EXPLORATION OF THE SEA
CIEM CONSEIL INTERNATIONAL POUR L'EXPLORATION DE LA MER

International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

H.C. Andersens Boulevard 44-46
DK-1553 Copenhagen V
Denmark
Telephone (+45) 33 38 67 00
Telefax (+45) 33 93 42 15
www.ices.dk
info@ices.dk

ISSN number: 2618-1371

ICES Scientific Reports

Volume 6 | Issue 99

WORKING GROUP ON MULTISPECIES ASSESSMENT METHODS (WGSAM)

Recommended format for purpose of citation:

ICES. 2024. Working Group on Multispecies Assessment Methods (WGSAM).
ICES Scientific Reports. 6:99. 77 pp. https://doi.org/10.17895/ices.pub.27919473

Editors

Michael A. Spence

Authors

Alan Baudron • Christopher Griffiths • Floor Soudijn • Francisco De Castro • Gustav Delius • Howard Townsend • Markus Varlund Strange • Michael A. Spence • Michael Thomson • Morten Vinther • Niamh Esmonde • Nis Sand Jacobsen • Pierre Yves Hernvann • Robert Thorpe • Sarah Gaichas • Thomas Bartos • Tyler Eddy • Valerio Bartolino • Vidette McGregor

ICES CIEM
International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

Contents

i Executive summary

The Working Group on Multispecies Assessment Methods (WGSAM) aims to advance the use of knowledge on predator-prey interactions for fisheries and ecosystem management. This report summarises the achievements of a 3-year cycle, focusing on the evaluation of key-runs and the skill assessment of multispecies models.

WGSAM released key-run models for various ICES regions, including the North Sea, Baltic Sea, and Georges Bank. The models for the North Sea and Baltic Sea provided updated estimates of predation mortality for key commercial stocks, which have been integrated into stock assessments.

Methods to assess the skill of multispecies models were established, focusing on understanding model prediction differences, choosing appropriate criteria, and comparing modelled dynamics with true dynamics. A publicly available simulated dataset was developed for skill assessment and model comparison.

Approaches for generating advice by comparing and combining multiple models were evaluated, including the R-package “EcoEnsemble” and methods for combining models with multiple drivers. Further developments included the "FishStomachs" R-package for stomach contents data, the MFDB data-handling package, and the growth of multispecies and ecosystem models. Efforts also focused on the impact of temperature on marine ecosystems, linking fisheries management with Good Environmental Status, developing new foodweb indicators, and assessing management goals and harvest control rules.

ii Expert group information

Expert group name	Working group on Multispecies Assessment Methods (WGSAM)
Expert group cycle	Multiannual
Year cycle started	2022
Reporting year in cycle	3/3
Chair	Michael A. Spence, UK
Meeting venues and dates	10-14 October 2022, Woods Hole, USA (>40 participants; 40% physical attendees and 60% remote)
	9-13 October 2023, Edinburgh, UK (9 physical attendees and approx. 15 remote combined)
	7-10 October 2024, Belfast, UK, (8 physical attendees and 18 in total)

1 List of terms of reference

ToR A. Regional updates: Review further progress and deliver key updates on multispecies modelling and ecosystem data analysis contributing to modelling throughout the ICES region

ToR B. Key-runs: Parametrisation of multispecies and ecosystem key-run models for different ICES regions. This includes standard update (limited to inclusion of recent data), extensive update (incl. new data and processes), and new key- runs.

ToR C. Skill assessment: Establish and apply methods to assess the skill of multispecies models intended for operational advice.

ToR D. Multi-model advice: Evaluate methods for generating advice by comparing and/or combining multiple models.

ToR E. MSE: Evaluate methods and applications for multispecies and ecosystem advice, including evaluation of management procedures and estimation of biological reference points under the uncertainties of climate change.

2 Summary of achievements during 2022–2024

- Keyruns: Modelling output and advisory products
  - SMS NS (see ICES 2024)
  - SMS Baltic (see ICES 2023)
  - Review of three models in Georges Bank (see ICES 2023)
- M values from SMS keyruns were used in the assessments of stocks in:
  - Cod (cod.27.46a7d20), haddock (had.27.46a20), herring (her.27.3a47d), sandeel (san.sa.1r, san.sa.2r, san.sa.3r, san.sa.4), sprat (spr.27.3a4), and whiting (whg.27.47d)
  - Baltic Sea herring (her.27.25-2932) and sprat (spr.27.22-32)
- Software and libraries
  - New R-Package "FishStomachs" for compilation of stomach contents data for the Baltic and North Seas which can be used to estimate population-level diets and foodweb rations. https://github.com/MortenVinther/FishStomachs
  - The "MFDB" (https://gadget-framework.github.io/mfdb/) data-handling package has been written in R, and is available on CRAN. It offers a "pre-canned" database schema suitable for storing data before inclusion into assessment models. It is not intended to be a primary source of data, rather a "staging area" for data to be gathered before any transformations are done to make the data suitable for e.g. a Gadget model.
  - "EcoEnsemble" – A general framework for combining ecosystem models in R.
  - R packages `atlantisom` for extracting simulated datasets from an Atlantis model run and `mskeyrun` a data package which stores multispecies simulated datasets (as well as real Georges Bank data used in the Georges Bank keyrun, see ToR b).
- Methodological developments:
  - Planned skill assessment for multispecies models using a common dataset with known dynamics, allowing both individual skill assessment and potential to assess ensemble skill. Initial fitting to multiple models (CEATTLE, State Space, Gadget, Hydra). Preliminary results: similarity of results across CEATTLE, state-space models for cod.
  - Extension of the ensemble model to account for models with similar drivers
- Collaborative paper Kempf, A., Spence, M. A., Lehuta, S., Trijoulet, V., Bartolino, V., Villanueva, M. C., and Gaichas, S. K. 2023. Skill assessment of models relevant for the implementation of ecosystem-based fisheries management. Fisheries Research, 268: 106845.
- A Euromarine workshop entitled "Multi-Modelled Marine Ecosystems (M3E)" was held in Nantes, France 11th-14th Oct. 2021.
- Papers associated to WGSAM activities are reported in the section "Relevant papers" under each ToR.

3 Final report on ToRs

3.1 ToR A. Regional updates: Review further progress and deliver key updates on multispecies modelling and ecosystem data analysis contributing to modelling throughout the ICES region

Over the course of 2022–2024, WGSAM received updates on modelling in the following regions: North Sea, Baltic Sea, US Northwest Atlantic Shelf, Tasman Sea, Gulf of Maine, Gulf of Alaska, Lake Ontario, US West Coast, Irish Sea, Celtic Sea and Mid Atlantic Bight.

Notable key points of the progress of multispecies and ecosystem modelling throughout the ICES regions are:

1) The number of multispecies and ecosystem models developed in connection with WGSAM is growing further, including models that have undergone in-depth reviews as keyruns and which are being maintained and regularly updated. For example, in this period, we used StrathE2E2 in the North Sea for the first time.
2) We are also increasingly using ensembles of models (e.g. EcoPath with EcoSim (EwE), LeMaRns, mizer, FishSUMS in the North Sea) built on the same ecosystems to support other ToRs and contribute to: (i) improve our understanding of the different modelling frameworks, (ii) build confidence on the outcomes of multispecies and ecosystem models (iii) better characterisation of uncertainties and areas where process understanding is weak.
3) The use of R-packages (e.g. Rceattle, Rpath, mizer, StrathE2E2, MFDB) for ecosystem modelling is continuing to progress and provide multi-species models and food-web models.
4) There is increased consideration for the use of fish stomach data to assess the diet or food ration (biomass eaten) of a given species within a given area to inform multi-species assessment models (e.g. SMS or Gadget). An R-package called 'FishStomach' has been developed to help calculate observed stomach data into population diet and biomass eaten for multispecies models.
5) Continued development on studying the effect of temperature to assess the impact of climate change on marine ecosystems (e.g. productivity, functional groups, food-web responses), particularly in the US and NZ.
6) There have also been advances in food-web modelling in the southern North Sea using EwE and Ecospace, which integrate new functional groups, improve parametrization and project food-web trajectories under different warming regimes.
7) Modelling has been used to link fisheries management with achievement of Good Environmental Status and hence the OSPAR process and wider ecosystem outcomes through an analysis of recovery timescales in the North Sea.
8) Models are continuing to contribute to ensembles, where there has been substantial methodological progress under ToR D. This is expected to lead to better predictions with more accurate characterisation of uncertainty.
9) Analysis is being undertaken in the US of new foodweb indicators which explore trade-offs between resilience and efficiency.

10) Information from generalist predators is being used to craft a "forage food index" for use in US East coast fisheries.
11) A modelling review of French capability revealed a large number of models, excellent analysis of current fleet segments and a high concentration of modelling expertise, but poor understanding of bottom-up processes including zooplankton. While half of the models had climate-ready functionality, there was poor understanding as to how fleets might adapt in the future and little idea as to what say 2050 might look like.
12) Work is ongoing to include seals within the mizer model for the North Sea. This is important because numbers are increasing and there is debate about whether their impact is significant at the fisheries or foodweb level.

Summaries of all presentations and discussions can be found in Annex 3 of this report.

Relevant Papers

Adams, Grant D., Kirstin K. Holsman, Steven J. Barbeaux, Martin W. Dorn, James N. Ianelli, Ingrid Spies, Ian J. Stewart, and André E. Punt. 2022. "An Ensemble Approach to Understand Predation Mortality for Groundfish in the Gulf of Alaska." *Fisheries Research* 251 (July): 106303. https://doi.org/10.1016/j.fishres.2022.106303.

Del Santo O'Neill, TJ, 2024. An efficient tool to find multispecies MSY for interacting fish stocks. Fish and Fisheries, 25(1), DOI: 10.1111/faf.12817

Fitzpatrick, Kimberly B., Brian C. Weidel, Michael J. Connerton, Jana R. Lantry, Jeremy P. Holden, Michael J. Yuille, Brian Lantry, *et al.* 2022. "Balancing Prey Availability and Predator Consumption: A Multispecies Stock Assessment for Lake Ontario." *Canadian Journal of Fisheries and Aquatic Sciences* 79 (9): 1529–45. https://doi.org/10.1139/cjfas-2021-0126.

Murray, DS. *et al.* 2024. Emerging issues in fisheries science by fisheries scientists, Journal of Fish Biology, 105(2). DOI: 10.1111/jfb.15683

Thorpe, RB. 2024. We need to talk about the role of zooplankton in marine food webs, Journal of Fish Biology, 105(2), DOI: 10.1111.jfb.15773

Thorpe, RB *et al.*, 2023. Can we use recovery timescales to define Good Environmental Status, Ecological Indicators, 155(3): 1470-160. DOI: 10.1016/j.ecolind.2023.110984

Thorpe, RB *et al.* 2022. The response of North Sea Ecosystem Functional Groups to Warming and Changes in Fishing, Frontiers in Marine Science, 9, DOI: 10.3389/fmars.2022.841909

3.2 ToR B. Key-runs: Parametrisation of multispecies and ecosystem key-run models for different ICES regions. This includes standard update (limited to inclusion of recent data), extensive update (incl. new data and processes), and new key-runs

WGSAM has a long-term commitment to advance the operational use of knowledge on species interactions (i.e., foodweb interactions, technical interactions) for advice in fisheries and ecosystem management. In the early 2010s (ICES 2013), WGSAM developed the concept of multispecies keyruns as quality assured model runs suitable to contribute to specific aspects of the ICES advice. For instance, multispecies keyruns are used to deliver natural mortality estimates for single species assessments of important North Sea and Baltic fish stocks.

3.2.1 Multispecies Model Review Criteria

For evaluation of the key runs the review criteria previously developed and adopted by WGSAM was used (ICES 2019, Kempf *et al.* 2023).For a full description of the review criteria see https://ices-eg.github.io/wg_WGSAM/ReviewCriteria.html and ICES 2021 The use of this protocol was useful and helped to standardise the reviews and ensure confidence that the review process is comprehensive and reliable.

The review criteria consist of the following six parts:

1. Is the model appropriate for the problem?
2. Is the scientific basis of the model sound?
3. Is the input data quality and parametrization sufficient for the problem?
4. Does model output compare well with observations?
5. Uncertainty
6. Previous Peer Review

For the Baltic Sea SMS key run (ICES 2023) and North Sea SMS key run (ICES 2024) the review criteria were implemented in full.

3.2.2 Baltic Sea

The Baltic Sea key run was performed in 2022 to review the estimates of natural mortality for the Central Baltic herring and Baltic sprat stocks. It explores the model's applicability, scientific basis, data quality, assumptions, and comparison with observations. The model was appropriate for estimating natural mortalities from 1974–2021, using predation by Eastern Baltic cod as the main factor. However, the model's dependency on Eastern Baltic cod data, which is assumed error-free, raises concerns about underestimating uncertainties.

The SMS model has been used in Baltic and North Sea fisheries assessments. Input data includes catch-at-age data, weight estimates, and cod stomach data, the latter being extensive but geographically limited. The model relies on certain assumptions, like constant predator-prey vulnerabilities and the availability of "Other Food," which might affect predictions.

While the model fits catch and survey data well for herring, sprat predictions show some discrepancies, especially in earlier years. The model's uncertainty estimates are limited, as they do not factor in potential errors in input data, particularly cod stock numbers.

Key Run Changes that were implemented in the 2022 keyrun:

- Stomach Data Compilation: The 2022 key run used the ‘FishStomach’ R package for compiling stomach data, introducing minimal differences in relative stomach contents compared to the 2019 key run.
- Food Ration Protocol: Revised food rationing protocol, splitting average quarterly consumption rates by quarters, leading to changes in M2 estimates.
- Fishing Mortality Configuration: Updated configuration to better capture temporal trends in age-specific fisheries effort, resulting in modest changes in M2 estimates.
- Cod as External Predator: Cod numbers are input to the model and assumed known without error, due to issues with age readings in Baltic cod.
- Alternative assumptions on cod consumption and background mortality (M1) were applied to reflect some of the structural uncertainties. Following request by

WKBALTPEL, alternative M estimates were derived as part of the keyrun and provided to the ensemble single species stock assessment of herring

Key recommendations include analyzing cod diet dynamics, splitting the "Other Food" category, and including spatial dynamics. Additionally, improvements are needed in estimating age-0 prey species and accounting for uncertainty in cod numbers. The review suggests the model could benefit from estimating multispecies FMSY values to aid fisheries management.

Where to find Input and output, the Baltic SMS keyrun

The description of model input data and configuration for the key-run can be found at (https://github.com/ices-eg/wg_WGSAM/tree/master/StockAnnex/Baltic/ StockAnnex_ICES_EB_SMS_2022_Configuration.pdf). An overview of input data and results can be found at (https://github.com/ices-eg/wg_WGSAM/tree/master/Baltic-2022-keyRun/HTML/ Baltic-2022-keyRun.html)

3.2.3 North Sea

North Sea SMS

At the WGSAM in 2023 the key-run for the North Sea using SMS was reviewed. SMS (Stochastic multispecies model, Lewy and Vinther 2004) model provides natural mortality estimates by age and year as input to single species stock assessments. The 2023 key-run provides natural mortality estimates for the assessments of cod, haddock, herring, Norway pout, southern North Sea sandeel, northern North Sea sandeel, sprat, and whiting. Natural mortality estimates are only used as input for the historic part of single species models and no forecast is needed. M estimates by age and quarter are a direct output of the SMS model. However, an assumption is needed for residual mortalities M1 while the predation mortalities M2 are estimated (M = M1+M2). The model provides estimates for the years 1974 to 2022.

Predators include both ICES assessed species (i.e., cod, haddock, saithe, whiting, mackerel) and species with given input population size (North Sea horse mackerel, western horse mackerel, grey gurnard, starry ray, hake, fulmar, gannet, great black backed gull, guillemot, herring gull, kittiwake, puffin, razorbill, grey seal, harbour porpoise). The assessed predators are parametrised using a combination of commercial and survey data (i.e., same input as for the single species assessments) except saithe and mackerel which are closely tuned to the ICES stock assessment by using number-at-age from the single species assessment models as input of SMS.

Overall, the model is considered appropriate to provide information on natural mortalities as input for the assessments. The 2023 North Sea SMS key-run is primarily an update of the 2020 key-run by: 1) extension of the input data and their update when the single species stock assessment input data were revised through benchmarks or inter-benchmarks, 2) complete revision of the grey seal and birds abundance time-series, 3) update of the diet calculation from the stomach data. Overall, the model structure and main assumptions are consistent with the previous key-run. WGSAM concluded that the model remains appropriate in relation to the purpose of providing predation mortality estimates. WGSAM recommends using these values as input to single species stock assessments.

A number of recommendations steam from this review:

- Recent samples of stomach contents should ideally be used in future SMS key-runs (including both the Baltic and North Sea), however new methods need to be developed to harmonise the new samples with those that occurred in the 1980s and in the Year of the Stomach (1991) which are still the base for this key-run. Potential solutions could be (i) to redefine the likelihood function to use individual observations, rather than the average diet, (ii) to work towards multiarea key-run models that would be inherently better to handle data spatial fragmentation in the sampling, or (iii) using statistical models to standardize the stomach data for changes in the spatial distribution, predator size, quarter, or other sampling variables. Currently, the new stomach database is still continuously being updated with new and old data collected over the years, and as the number of samples among years increase, this should represent a strong incentive to use the full time-series of samples for future key-runs.

- From the review it emerged that it would be beneficial that modellers in WGSAM are better supported by the stock assessment working groups in the preparation, maintenance and documentation of the time-series input for the North Sea key-run. That would help to consolidate the input data and would provide the relevant information to interpret changes in between key-run well in advance of the working group meeting and key-run review.
- Similarly, the review suggests that update of time-series for birds would largely benefit from contribution by bird experts (i.e., JWGBIRD).

Where to find Input and output, the SMS program and R-scripts

The description of model input data and configuration (the stock annex) for the key-run can be found at (insert link to ICES stock Annex database where we hope that ICES will include the file https://github.com/ices-eg/wg_WGSAM/tree/master/StockAnnex/NorthSea/2023_NorthSea.pdf). An overview of input data and results can be found at (link to ICES library where ICES will insert the file https://github.com/ices-eg/wg_WGSAM/tree/master/NorthSeaKeyRun_2023/HTML/NS_2023_key_run.html.)

The Github for WGSAM (https://github.com/ices-eg/wg_WGSAM/tree/master) includes several directories from this and previous key-runs. The most relevant directories for the most recent North Sea key-run are:

- NortSeaKeyRun_2023: Input and output from the SMS North Sea key run made at the 2023 WGSAM
- SMS_ADMB: AD Model Builder source code for the SMS North Sea key run
- SMS_R_prog: R scripts for preparing, running and presenting results from an SMS run
- SMS_Stomachs: R scripts for compilation of stomach contents observations into population diet

3.2.4 Georges Bank

At the 2022 meeting, WGSAM reviewed a suit of models to address long-term multiple management objectives in the context of ecosystem-based fisheries management for the Georges Bank, namely a multispecies production model (MSSPM), a multispecies length-based model (Hydra), and a mass-balanced ecosystem model (Rpath). While none of the three model implementations was sufficiently mature as a keyrun at the time of the review, the WGSAM's evaluation was largely based on the review criteria for model keyruns (ICES 2021) which (i) offered an excellent opportunity to test the evaluation criteria on largely different modelling frameworks and (ii) provided a good feedback for further work on those models and their possible future eligibility as keyruns. Main outcomes from the WGSAM's review of the Georges Bank data and modelling suits:

- The Georges Bank is an ecosystem production unit (EPU) but its secondary production is influenced by strong connectivity with adjacent ecosystems (e.g., presence of ecologically and economically relevant species with larger distributions and migratory species transient within the EPU) which may require a broader perspective across the Northeast continental shelf.
- A set of relevant R-packages were developed to prepare commercial (comlandr, mscatch) and survey (survdat) data to enhance transparency and reproducibility.
- MSSPM is able to a handle relatively sparse and shortage of data, which is not necessarily the case for the Georges Bank, but it is a relatively simple multispecies model that can provide preliminary estimates of species interdependent dynamics and carrying capacity by species, trophic guild and entire ecosystem.
- Hydra is a length-structured model. Latest developments of the model from a simulation to an estimation tool able to fit multiple datasets (incl. stomach data) give it the potentiality to be used in the future as a multispecies assessment model.
- Rpath implementation is intended to describe the structure and flow of energy through the ecosystem, by quantifying the food web interactions for a high number of species or trophic groups. The model was found potentially suitable to assess the consequences of certain management strategies from a broad ecosystem perspective, hence it has a scope for MSE when used in conjunction with other multispecies models. Rpath implementations for adjacent areas, such as the Gulf of Maine and the Middle Atlantic Bight, already exist and their linking could be potentially important to address relevant energy flows across ecosystems.

3.3 ToR C. Skill assessment: Establish and apply methods to assess the skill of multispecies models intended for operational advice

Multispecies and ecosystem models are complex, which can lead to high uncertainty in predictions (Hill *et al.* 2007). Skill assessment compares different model predictions with the truth from a system (Stow *et al.* 2009; Olsen *et al.* 2016). Understanding model skill will help develop better models, as well as understand which models are most effective in which management situations. This work was aimed at assessing the performance of models intended for strategic or tactical management advice. Evaluation required work towards standardisation for cross -model comparison. This ToR also dealt with evaluation of methods for model calibration and data weighting in the context of multispecies modelling.

WGSAM made considerable progress on skill assessment methods for multispecies models during 2022–2024. The group published a paper (Kempf *et al.* 2023) recommending best practices for model skill assessment. The group also developed open access tools and datasets for skill assessment and initiated an ambitious project to evaluate skill across several differently structured multispecies models. Summaries of all presentations and discussions can be found in Annex 4 of this report.

The main themes and outcomes of skill assessment investigations were:

- **Predictions from different models can differ significantly, and it is important to understand why.** For instance, Robert Thorpe presented work carried out with Gustav Delius, Michael Spence, and Georg Engelhard on methods to evaluate and compare the predictions and robustness of mizer and EwE and to understand why mizer is so much more robust to high levels of fishing than EwE, and take steps towards understanding whether predicted outcomes are realistic. Both models can produce a skillful hindcast, despite having very different robustness to high levels of fishing, so additional sources of information are needed to discriminate between them.
- **Precautions should be taken in choosing the criterion for assessing model skill.** Vanessa Trijoulet demonstrated that common standardized residuals cannot be used for compositional data model validation and can lead to wrong conclusions due to correlation in the observations that are propagated into the residuals. One-step-ahead (OSA) residuals, on the other hand, have the correct properties when the model is correct: independent, normally distributed, with mean zero, and with variance one. An R-package "compResidual" and a Template Model Builder (TMB) are available that allow estimation of the OSA residuals externally to the model and inside the model. These developments are relevant to any type of compositional data for single and multi-species models, i.e., aged- and length-structured data, stomach content data.
- **Skill assessment requires comparison of modelled with true dynamics rather than simply fit to observations.** WGSAM used Atlantis ecosystem model output to generate input datasets for other multispecies models, which allows evaluation of performance against true (simulated) ecosystem dynamics. Atlantis models can be run using different climate forcing, fishing, and other scenarios. Users of package atlantisom will be able to reproducibly specify fishery independent and fishery dependent sampling in space and time, as well as species-specific catchability, selectivity, and other observation processes for any Atlantis scenario.
- **A publicly available simulated dataset (mskeyrun) was developed for skill assessment using the atlantisom package.** The dataset includes indices of abundance and catch as well as age, length, and diet composition data for 11 species.
- **A cross model comparison study was designed and is currently being implemented (Figure 3.3.1).** To date, 4 models (Gadget2, CEATTLE, StateSpace, and Hydra) have been fit to the simulated data for at least one and up to 11 species within different age and length structured model frameworks. Initial parameterizations have been started for mizer, LeMans, and Rpath.

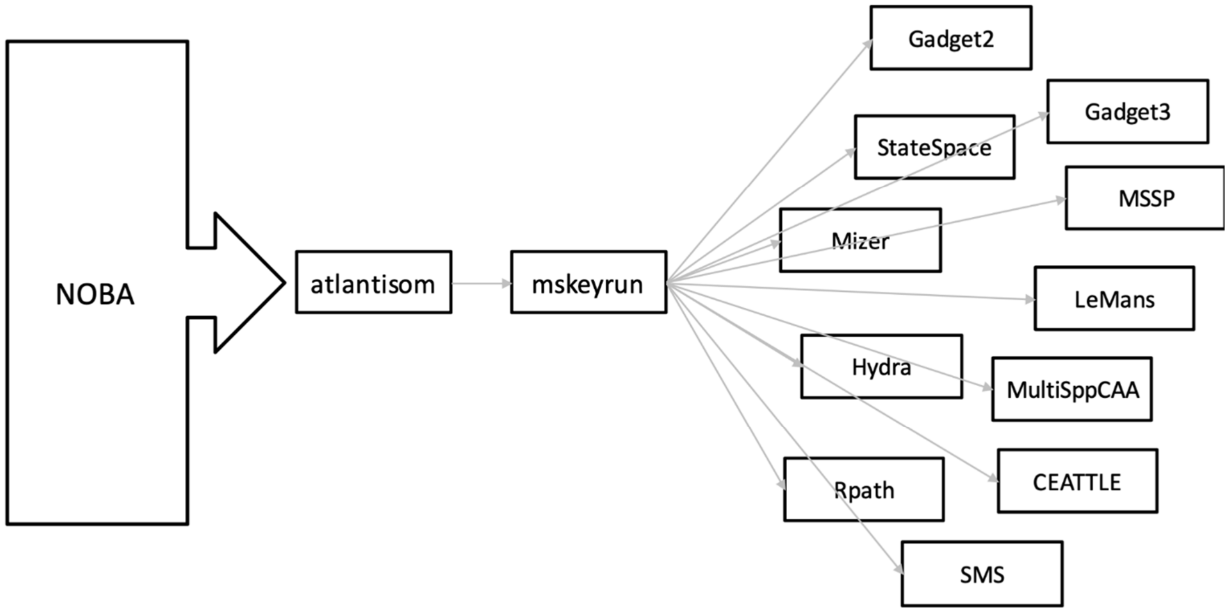


Figure 3.3.1. Model skill assessment design. The Norwegian-Barents (NOBA, (Hansen *et al.* 2016, 2019)) Atlantis model output from a climate-neutral run is used by `atlantisom` package functions to create the simulated datasets in the `mskeyrun` package. Up to 11 multispecies and ecosystem models are using the mskeyrun simulated data as inputs.

The publication of a best practices paper and a simulated dataset with sufficient complexity for multispecies model skill assessment are key outcomes for 2022–2024, in addition to the development of open source methods for deriving simulated datasets. While multiple models have been preliminarily fit to the simulated data, more work remains to fully compare the skill of these models and to develop guidelines for fitting models to data based on the skill assessment project. Therefore, it is recommended that a skill assessment ToR be continued for the next three years.

Relevant papers

Kempf, A., Spence, M. A., Lehuta, S., Trijoulet, V., Bartolino, V., Villanueva, M. C., and Gaichas, S. K. 2023. Skill assessment of models relevant for the implementation of ecosystem-based fisheries management. Fisheries Research, 268: 106845.

Trijoulet, V., Albertsen, C. M., Kristensen, K., Legault, C. M., Miller, T. J., and Nielsen, A. 2023. Model validation for compositional data in stock assessment models: Calculating residuals with correct properties. Fisheries Research, 257: 106487.

3.4 ToR D. Multi-model advice: Evaluate methods for generating advice by comparing and/or combining multiple models

The R-package EcoEnsemble that allows operational use for Spence *et al.* 2018 has been published. It has been used to explore concepts of multispecies precautionary and multispecies maximum sustainable yield in the North Sea. WGSAM also explored the sensitivity of this ensemble with the prior distribution of the parameters. The ensemble model of Spence *et al.* 2018 has been extended to account for different drivers, e.g. inputs from different Earth system models (ESMs). This work was used in a study of climate effects in on sardines in the California Current. Another extension combines single-species and multispecies models to synthesis forecasts seamlessly across time scales.

Fishing mortality rates from the whole of the 20th century were estimated from two multispecies models and combined using an ensemble model. The study provides valuable insights into short-term dynamics in the periods after World Wars one and two, and a timeseries of fishing mortality that can be used to drive other models.

As part of the Fisheries and Marine Ecosystem Model Intercomparison Project (FishMIP), different ecosystem models, including a variety of approaches, were driven with climate forcings from ESMs on both a global and a regional scale to explore the consequences of climate change on the marine ecosystem.

The "multi-facet approach" was developed. This approach involves using multiple models to inform different facets of the same question, rather than trying to implement the same scenario across all models. The approach can help address common issues in multi-model exercises, such as narrow scope, little operational use of scenarios, and inappropriate use of models, by focusing on the ecological meaning of the chosen implementation and the similarities and differences between models.

Summaries of all presentations and discussions can be found in Annex 5 of this report.

Relevant papers

Spence, M. A., Martindale, J. A., & Thomson, M. J. (2023). EcoEnsemble: A general framework for combining ecosystem models in R. Methods in Ecology and Evolution, 14, 2011–2018. https://doi.org/10.1111/2041-210X.14148

Spence, M. A., Martindale, J. A., Alliji, K., Bannister, H. J., Thorpe, R. B., Walker, N. D., Mitchell, P. D., Kerr, M. R., Dolder P. J., (2024). Assessing the effect of multispecies interactions on precautionary reference points using an ensemble modelling approach: A North Sea case study, Fisheries Research, 280, 107160.

Ruiz-Diaz R, Pennino MG, Fisher JAD, Eddy TD (2024) Decadal changes in biomass and distribution of key fisheries species on Newfoundland's Grand Banks. PLOS ONE 19(4): e0300311. https://doi.org/10.1371/journal.pone.0300311

3.5 ToR E. MSE: Evaluate methods and applications for multispecies and ecosystem advice, including evaluation of management procedures and estimation of biological reference points under the uncertainties of climate change

At the core of WGSAM lies the idea that the dynamics of a fish stock depend on other fish stocks. Therefore, the application of MSYs to single fish stocks is considered an unsuitable management principle. For example, in a predator-prey system, maximising the yield of the predator requires its prey to be unexploited whereas by harvesting the predator the yield of the prey can be maximised (Soudijn *et al.* 2021).

Within ToR E, WGSAM aims to consider methods to estimate multispecies rather than single species management targets, such as for example multispecies MSY and FMSY values, or multispecies biological reference points, which can be used to assess the ecosystem state. Advances of multispecies modelling in the context of fisheries advice has stimulated a new discussion which has also benefited from the increasing interest and use of management strategy evaluation approaches (MSE and MSE-like). The complexity of multispecies models, their ability to represent ecosystems components beyond the exploited stocks, and the different implementation that they offer of ecological processes in comparison to stock assessment models (especially of end-to-end models), makes them useable tools for the development of operating models. However, these advantages come at the expense of complexity in their implementation within an MSE. Several contributions have been presented that consider MSE in multispecies approaches during this 3-years cycle for which the main outcomes are summarised below:

- A projection study with the multispecies modelling framework GADGET (Globally applicable Area Disaggregated General Ecosystem Toolbox) of the Baltic Sea evaluated how the current F_{MSY} management framework responds to climate-related shocks. Climate-related changes in the recruitment of Cod, Herring and Sprat were simulated using coupled physical-biogeochemical models and Gadget-SSB recruitment estimates for 2021–2030. Subsequently, fish stocks' responses to recruitment shocks (for example due to heatwaves) were evaluated using metrics for resistance (i.e. ability of the stocks to oppose perturbations) and resilience (i.e. ability of the stocks to recover from perturbations). It was found that a precautionary fishing level at the lower limit of the F_{MSY} range was needed to create a buffer against the effect of shocks, particularly for herring and less so for sprat.
- The possibility of including Management Strategy Evaluation in an Ecosystem-Based fisheries management approach was tested by combining Hydra, a length-based multispecies, multifleet, spatial model (Gaichas *et al.*, 2017) and a groundfish MSE framework of New England (Mazur *et al.*, 2023). The process included iterative communication with the management committee. When complex stock management (based on multiple species) was included in the models, catches were found to be higher and less variable without harming any of the stocks' conservation.
- The performance of management goals from single species assessments was tested with CEATTLE (Climate-Enhanced, Age-based Model, with Temperature-Based, Trophic Linkages and Energenics) (Holsman *et al.* 2016) in the Eastern Bering Sea and Gulf of Alaska. The effect of climate change was considered in the models (https://github.com/grant-dadams/Rceattle, Punt *et al.* 2024).
- The performance of several harvest control rules for the Flemish Cap was evaluated using GADGET (Globally applicable Area Disaggregated General Ecosystem Toolbox), a

multispecies model of a group of interacting commercial stocks (Perez-Rodrigues *et al.* 2017). Due to the strong interactions in the model, it was not possible to achieve precautionary exploitation of all the stocks at the same time for any of the harvest control schemes that were simulated (Perez-Rodrigues *et al.* 2022).

- The 12 alternative interpretations of the UK Common Fisheries Policy for the North Sea were explored using the ensemble model described by Spence *et al.* (2024). The findings indicated that MSY estimates based on a single model perform poorly and a fully multispecies interpretation is expected to perform best in the long term, although challenges remain in gaining stakeholder acceptance for this approach.
- The potential for applying precautionary management reference points based on multispecies assessment models was explored for nine stocks in the North Sea. An ensemble model analysis was conducted with a range of multispecies models of the North Sea. While the robustness to fishing was found to be different per model, no fishing strategy was found to be precautionary for all stocks simultaneously (Spence *et al.* 2024). Even no fishing at all was only considered precautionary for six out of nine stocks based on current precautionary management reference points, which are based on single-species stock assessments. The study concludes that new methods are needed to determine precautionary reference points with a multispecies perspective (Spence *et al.* 2024).
- A new multispecies indicator of biological risk was proposed (Murray *et al.* 2024). The indicator (the relative risk of depletion) is designed to provide a holistic and scalable approach to assessing the risk of stock depletion across multiple species within an ecosystem. The key principles of the new indicator involve the ecosystem components being equally weighted, the status of each component being expressed relative to its unfished state (ranging from 0 to 1), making the rate of change of risk of a component inversely proportional to its status, and adding a societal tolerance contribution. The new indicator has several advantages over previous metrics, such as being scalable by ecosystem size, applicable to all ecosystem components, and ensuring that no approach to extinction events is deemed acceptable. It also retains the need for stakeholder input to determine acceptable risk levels, aligning with the Common Fisheries Policy (CFP) and the Marine Strategy Framework Directive (MSFD).
- A portfolio optimization approach (Brewster *et al.* 2023a) was applied to six US fisheries regions (Alaska, Gulf of Maine, Mid-Atlantic, New England, West Coast and South Atlantic). This empirical approach was based on historical patterns in landings and revenue data. A comparison was made between a portfolio optimization based on single-species and ecosystem-level fisheries management objectives, by respectively ignoring or including correlations in species revenues (Townsend *et al.* 2024). The analysis indicated that fisheries management strategies based on the ecosystem management objectives led to a lower risk in revenues than fisheries management strategies based on single-species management objectives. It should be noted that the optimizations in this approach are based on economic risks and do not directly take ecological risk into consideration.
- The Common sole (*Solea solea*) in the Bay of Biscay faces challenges in sustainable exploitation due to mixed demersal fisheries and changes in productivity. At the WGSAM 2024, a management strategy evaluation (MSE) framework was introduced, using the ISIS-Fish model to simulate fisheries dynamics without representing species interactions. This framework connects a sole stock assessment model to ISIS-Fish, incorporating fisheries, population and management modules. It evaluates the impact of mixed fisheries, environmental drivers, and management procedures on sole productivity. The framework uses a nursery habitat suitability model with river flow as a covariate to test alternative management procedures. Preliminary results show that environmentally-informed harvest control rules contribute to stock rebuilding but increase catch variability.

The current single-species fisheries management strategies seem to not always hold up when they are considered in a multispecies framework. Two of the studies mentioned above (Spence *et al.* 2024, Perez-Rodrigez *et al.* 2022) have indicated that current precautionary biological reference points, which have been defined based on single species models, are simply impossible to reach when they are considered in multispecies models. This illustrates the importance of considering the interrelations between different stocks when defining biological reference points. If not, fisheries management objectives may be simply unattainable in the real world. There is still quite some work to be done to come up with multispecies biological reference points that are deemed acceptable and sustainable from an ecological, economic and a social point of view.

Summaries of all presentations and discussions can be found in Annex 6 of this report.

Relevant papers

Mazur, M. D. *et al.* (2023). "Consequences of ignoring climate impacts on New England groundfish stock assessment and management". In: Fisheries Research 262, p. 106652.

Murray, D. S., Campón-Linares, V., O'Brien, C. M., Thorpe, R. B., Vieira, R. P., & Gilmour, F. (2024). Emerging issues in fisheries science by fisheries scientists. Journal of Fish Biology, 105(2), 557–563.

Pérez-Rodríguez A, Umar I, Goto D, Howell D, Mosqueira I, González-Troncoso D. 2022. Evaluation of harvest control rules for a group of interacting commercial stocks using a multispecies MSE framework. *Canadian Journal of Fisheries and Aquatic Sciences* **79**: 1302–1320.

Punt AE, Dalton MG, Adams GD, Barbeaux SJ, Cheng W, Hermann AJ, Holsman KK, Hulson P-JF, Hurst TP, Rovellini A. 2024. Capturing uncertainty when modelling environmental drivers of fish populations, with an illustrative application to Pacific Cod in the eastern Bering Sea. Fisheries Research 272: 106951.

Spence, M. A., Martindale, J. A., Alliji, K., Bannister, H. J., Thorpe, R. B., Walker, N. D., Mitchell, P. D., Kerr, M. R., Dolder P. J., (2024). Assessing the effect of multispecies interactions on precautionary reference points using an ensemble modelling approach: A North Sea case study, *Fisheries Research*, 280, 107160.

Townsend H, Link JS, DePiper G, Brewster LR, Cadrin SX, Edwards F. n.d. Multispecies Portfolios of U.S. Marine Fisheries: Ecosystem-Based Fisheries Management Reduces Economic Risk. Fisheries n/a DOI: 10.1002/fsh.11152.

3.6 References

Brewster L, Edwards F, Link J, Cadrin S. 2023. Using portfolio theory to improve the management of living marine resources: a demonstration for New England fisheries. Report to New England Fishery Management Council's Scientific and Statistical Committee.

Gaichas, S. K. *et al.* (2017). "Combining stock, multispecies, and ecosystem level fishery objectives within an operational management procedure: simulations to start the conversation". In: ICES Journal of Marine Science 74.2, pp. 552-565.

Hansen, Cecilie, Kenneth F. Drinkwater, Anne Jähkel, Elizabeth A. Fulton, Rebecca Gorton, and Mette Skern-Mauritzen. 2019. "Sensitivity of the Norwegian and Barents Sea Atlantis End-to-End Ecosystem Model to Parameter Perturbations of Key Species." Edited by Athanassios C. Tsikliras. PLOS ONE 14 (2): e0210419. https://doi.org/10.1371/journal.pone.0210419.

Hansen, Cecilie, Mette Skern-Mauritzen, Gro van der Meeren, Anne Jähkel, and Kenneth F. Drinkwater. 2016. Set-up of the Nordic and Barents Seas (NoBa) Atlantis Model. Fisken Og Havet;2-2016. Havforskningsinstituttet. https://imr.brage.unit.no/imr-xmlui/bitstream/handle/11250/2408609/FoH_2-2016.pdf?sequence=1&isAllowed=y.

Hill, Simeon L., George M. Watters, André E. Punt, Murdoch K. McAllister, Corinne Le Quéré, and John Turner. 2007. "Model Uncertainty in the Ecosystem Approach to Fisheries." Fish and Fisheries 8 (4): 315–36. https://doi.org/10.1111/j.1467-2979.2007.00257.x.

Holsman KK, Ianelli J, Aydin K, Punt AE, Moffitt EA. 2016. A comparison of fisheries biological reference points estimated from temperature-specific multi-species and single-species climate-enhanced stock assessment models. *Deep Sea Research Part II: Topical Studies in Oceanography* **134**: 360–378.

ICES. 2013. Interim Report of the Working Group on Multispecies Assessment Methods (WGSAM), 21-25 October 2013, Stockholm, Sweden. ICES CM 2013/SSGSUE:10. 99 pp. https://doi.org/10.17895/ices.pub.9092

ICES. 2019. Working Group on Multispecies Assessment Methods (WGSAM). ICES Scientific Reports. 1:91. 320 pp. http://doi.org/10.17895/ices.pub.5758

ICES. 2021. Working Group on Multispecies Assessment Methods (WGSAM; outputs from 2020 meeting). ICES Scientific Reports. 3:10. 231 pp. https://doi.org/10.17895/ices.pub.7695

ICES. 2023. Working Group on Multispecies Assessment Methods (WGSAM; outputs from 2022 meeting). ICES Scientific Reports. 5:12. 233 pp. https://doi.org/10.17895/ices.pub.22087292

ICES. 2024. Working Group on Multispecies Assessment Methods (WGSAM; outputs from 2023 meeting). ICES Scientific Reports. 6:13. 218 pp. https://doi.org/10.17895/ices.pub.25020968

Kempf, A., Spence, M. A., Lehuta, S., Trijoulet, V., Bartolino, V., Villanueva, M. C., and Gaichas, S. K. 2023. Skill assessment of models relevant for the implementation of ecosystem-based fisheries management. Fisheries Research, 268: 106845.

Lewy, P., and Vinther, M. 2004. A stochastic age-length-structured multi-species model applied to North Sea stocks. 20. 33 pp.

Mazur, M. D. *et al.* (2023). "Consequences of ignoring climate impacts on New England groundfish stock assessment and management". In: Fisheries Research 262, p. 106652.

Murray, D. S., Campón-Linares, V., O'Brien, C. M., Thorpe, R. B., Vieira, R. P., & Gilmour, F. (2024). Emerging issues in fisheries science by fisheries scientists. Journal of Fish Biology, 105(2), 557–563.

Olsen, Erik, Gavin Fay, Sarah Gaichas, Robert Gamble, Sean Lucey, and Jason S. Link. 2016. "Ecosystem Model Skill Assessment. Yes We Can!" Edited by Carlo Nike Bianchi. PLOS ONE 11 (1): e0146467. https://doi.org/10.1371/journal.pone.0146467.

Pérez-Rodríguez A, Howell D, Casas M, Saborido-Rey F, Ávila-de Melo A. 2017. Dynamic of the Flemish Cap commercial stocks: use of a Gadget multispecies model to determine the relevance and synergies among predation, recruitment, and fishing. *Canadian Journal of Fisheries and Aquatic Sciences* **74**: 582–597.

Pérez-Rodríguez A, Umar I, Goto D, Howell D, Mosqueira I, González-Troncoso D. 2022. Evaluation of harvest control rules for a group of interacting commercial stocks using a multispecies MSE framework. *Canadian Journal of Fisheries and Aquatic Sciences* **79**: 1302–1320.

Punt AE, Dalton MG, Adams GD, Barbeaux SJ, Cheng W, Hermann AJ, Holsman KK, Hulson P-JF, Hurst TP, Rovellini A. 2024. Capturing uncertainty when modelling environmental drivers of fish populations, with an illustrative application to Pacific Cod in the eastern Bering Sea. Fisheries Research 272: 106951.

Soudijn FH, Daniël van Denderen P, Heino M, Dieckmann U, de Roos AM. 2021. Harvesting forage fish can prevent fishing-induced population collapses of large piscivorous fish. *Proceedings of the National Academy of Sciences* **118**: e1917079118. DOI: 10.1073/pnas.1917079118

Spence, MA, Blanchard, JL, Rossberg, AG, *et al.* A general framework for combining ecosystem models. Fish Fish. 2018; 19: 1031– 1042. https://doi.org/10.1111/faf.12310

Spence, M. A., Martindale, J. A., Alliji, K., Bannister, H. J., Thorpe, R. B., Walker, N. D., Mitchell, P. J., Kerr, M. R., Dolder P. J., (2024). Assessing the effect of multispecies interactions on precautionary reference points using an ensemble modelling approach: A North Sea case study, Fisheries Research, 280, 107160.

Townsend H, Link JS, DePiper G, Brewster LR, Cadrin SX, Edwards F. n.d. Multispecies Portfolios of U.S. Marine Fisheries: Ecosystem-Based Fisheries Management Reduces Economic Risk. Fisheries n/a DOI: 10.1002/fsh.11152.

Stow, Craig A., Jason Jolliff, Dennis J. McGillicuddy, Scott C. Doney, J. Icarus Allen, Marjorie A. M. Friedrichs, Kenneth A. Rose, and Philip Wallhead. 2009. "Skill Assessment for Coupled Biological/Physical Models of Marine Systems." Journal of Marine Systems, Skill assessment for coupled biological/physical models of marine systems, 76 (1): 4–15. https://doi.org/10.1016/j.jmarsys.2008.03.011.

Annex 1: List of participants

WGSAM 2024 meeting

Name	Institute	Country	Email
Michael Spence (chair)	Centre for Environment, Fisheries and Aquaculture Science (CEFAS)	United Kingdom	michael.spence@cefas.gov.uk
Valerio Bartolino	Swedish University of Agricultural Sciences/ Institute of Marine Research	Sweden	valerio.bartolino@slu.se
Thomas Bartos	Centre for Environment, Fisheries and Aquaculture Science (CEFAS)	United Kingdom	Thomas.bartos@cefas.gov.uk
Alan Baudron	Marine Scotland Science	United Kingdom	Alan.Baudron@gov.scot
Francisco De Castro	Fisheries & Aquatic Ecosystems. AFBI	United Kingdom	francisco.decastro@afbini.gov.uk
Gustav Delius	Department of Mathematics, University of York	United Kingdom	gustav.delius@york.ac.uk
Tyler Eddy	Fisheries & Oceans Canada (DFO) and the Canadian Association of Prawn Producers & Northern Coalition	Canada	Tyler.Eddy@mun.ca
Niamh Esmonde	Fisheries & Aquatic Ecosystems. AFBI	United Kingdom	Niamh.esmonde@afbini.gov.uk
Sarah Gaichas	NOAA Northeast Fisheries Science Center, Woods Hole	USA	Sarah.Gaichas@noaa.gov
Christopher Griffiths	Swedish University of Agricultural Sciences/ Institute of Marine Research	Sweden	christopher.griffiths@slu.se
Pierre Yves Hernvann	IFREMER	France	pierre.yves.hernvann@gmail.com
Nis Sand Jacobsen	DTU-Aqua	Denmark	nsja@aqua.dtu.dk
Vidette McGregor	National Institute of Water and Atmospheric Research (NIWA)	New Zealand	Vidette.McGregor@niwa.co.nz
Floor Soudijn	Wageningen Marine Research	Netherlands	floor.soudijn@wur.nl
Markus Varlund Strange	DTU-Aqua	Denmark	mavast@aqua.dtu.dk
Michael Thomson	Centre for Environment, Fisheries and Aquaculture Science (CEFAS)	United Kingdom	Michael.thomson@cefas.gov.uk
Robert Thorpe	Centre for Environment, Fisheries and Aquaculture Science (CEFAS)	United Kingdom	Robert.thorpe@cefas.gov.uk
Howard Townsend	NOAA Fisheries	USA	Howard.Townsend@noaa.gov
Morten Vinther	DTU-Aqua	Denmark	mv@aqua.dtu.dk

WGSAM 2023 meeting

Name	Institute	Country (of institute)	Email
Michael Spence (chair)	Centre for Environment, Fisheries and Aquaculture Science (CEFAS)	United Kingdom	michael.spence@cefas.gov.uk
Valerio Bartolino (chair)	Swedish University of Agricultural Sciences/ Institute of Marine Research	Sweden	valerio.bartolino@slu.se
Grant Adams	School of Aquatic and Fishery Sciences, University of Washington	United States	adamsgd@uw.edu
Alan Baudron	Marine Scotland Science	United Kingdom	Alan.Baudron@gov.scot
Maria Ching Villanueva	IFREMER	France	Ching.villanueva@ifremer.fr
Thomas Del Santo O'Neill	Queen Mary, University of London	United Kingdom	delsantooneillthomas@gmail.com
Gustav Delius	Department of Mathematics, University of York	United Kingdom	gustav.delius@york.ac.uk
Sarah Gaichas	NOAA Northeast Fisheries Science Center, Woods Hole	USA	Sarah.Gaichas@noaa.gov
Christopher Griffiths	Swedish University of Agricultural Sciences/ Institute of Marine Research	Sweden	christopher.griffiths@slu.se
Nis Sand Jacobsen	DTU-Aqua	Denmark	nsja@aqua.dtu.dk
Alexander Kempf	Thuenen Institute of Sea Fisheries	Germany	alexander.kempf@thuenen.de
Sean Lucey	NOAA Northeast Fisheries Science Center, Woods Hole	USA	Sean.Lucey@noaa.gov
Vidette McGregor	National Institute of Water and Atmospheric Research (NIWA)	New Zealand	Vidette.McGregor@niwa.co.nz
Alfonso Perez	IEO	Spain	alfonso.perez@ieo.csic.es
Miriam Püts	Thuenen Institute of Sea Fisheries	Germany	miriam.puets@thuenen.de
Floor Soudijn	Wageningen Marine Research	Netherlands	floor.soudijn@wur.nl
Michael Thomson	Centre for Environment, Fisheries and Aquaculture Science (CEFAS)	United Kingdom	Michael.thomson@cefas.gov.uk
Robert Thorpe	Centre for Environment, Fisheries and Aquaculture Science (CEFAS)	United Kingdom	Robert.thorpe@cefas.gov.uk
Howard Townsend	NOAA Fisheries	USA	Howard.Townsend@noaa.gov
Vanessa Trijoulet	DTU-Aqua	Denmark	vtri@aqua.dtu.dk
Morten Vinther	DTU-Aqua	Denmark	mv@aqua.dtu.dk

WGSAM 2022 meeting

Name	Institute	Country (of institute)	Email
Valerio Bartolino (co-chair)	Swedish University of Agricultural Sciences/ Institute of Marine Research	Sweden	valerio.bartolino@slu.se
Michael Spence (co-chair)	Centre for Environment, Fisheries and Aquaculture Science (CEFAS)	United Kingdom	michael.spence@cefas.gov.uk
Sarah Gaichas	NOAA Northeast Fisheries Science Center, Woods Hole	USA	Sarah.Gaichas@noaa.gov
Sean Lucey	NOAA Northeast Fisheries Science Center, Woods Hole	USA	Sean.Lucey@noaa.gov
Vanessa Trijoulet	DTU-Aqua	Denmark	vtri@aqua.dtu.dk
Morten Vinter	DTU-Aqua	Denmark	mv@aqua.dtu.dk
Alfonso Perez	IEO	Spain	alfonso.perez@ieo.csic.es
Chris Legault	NOAA Northeast Fisheries Science Center, Woods Hole	USA	chris.legault@noaa.gov
Sigrid Lehuta	IFREMER	France	sigrid.lehuta@ifremer.fr
Kiersten Curti	NOAA Northeast Fisheries Science Center, Woods Hole	USA	kiersten.curti@noaa.gov
Gavin Fay	School for Marine Science and Technology, University of Massachusetts Dartmouth	USA	gfay@umassd.edu
Andy Beet	NOAA Northeast Fisheries Science Center, Woods Hole	USA	andrew.beet@noaa.gov
Howard Townsend	NOAA Fisheries, Silver Spring	USA	Howard.Townsend@noaa.gov
Robert Gamble	NOAA Northeast Fisheries Science Center, Woods Hole	USA	Robert.Gamble@noaa.gov
Ron Klasky	NOAA Fisheries, Silver Spring	USA	ronald.klasky@noaa.gov
Christina Perez	School for Marine Science and Technology, University of Massachusetts Dartmouth	USA	mperez12@umassd.edu
Maxwell Grezlik	School for Marine Science and Technology, University of Massachusetts Dartmouth	USA	mgrezlik@umassd.edu
Floor Soudijn	Wageningen Marine Research	Netherlands	floor.soudijn@wur.nl
Alan Baudron	Marine Scotland Science	United Kingdom	Alan.Baudron@gov.scot
Andrea Belgrano	Swedish University of Agricultural Sciences, Institute of Marine Research and Swedish Institute for the Marine Environment (SIME)	Sweden	andrea.belgrano@slu.se
Jamie Tam	Bedford Institute of Oceanography,· Oceans and Ecosystem Sciences Division	Canada	Jamie.Tam@dfo-mpo.gc.ca
Jacob Bentley	Natural England, London	United Kingdom	Jacob.Bentley@sams.ac.uk

Christopher Griffiths	Swedish University of Agricultural Sciences/ Institute of Marine Research	Sweden	christopher.griffiths@slu.se
Francisco De Castro	Fisheries & Aquatic Ecosystems (AFBI)	United Kingdom	francisco.decastro@afbini.gov.uk
Carolyn McKeon	Agri-Food and Biosciences Institute (AFBI)	United Kingdom	mckeonc2@tcd.ie
Miriam Puts	Thünen Institute of Sea Fisheries	Germany	miriam.puest@thuenen.de
Gustav Delius	Department of Mathematics, University of York	United Kingdom	gustav.delius@york.ac.uk
Michael Thomson	Centre for Environment, Fisheries and Aquaculture Science (CEFAS)	United Kingdom	michael.thomson@cefas.gov.uk
James Martindale	Centre for Environment, Fisheries and Aquaculture Science (CEFAS)	United Kingdom	james.martindale@cefas.co.uk
Alex Holdgate	Centre for Environment, Fisheries and Aquaculture Science (CEFAS)	United Kingdom	alex.holgate@cefas.gov.uk
Mikaela Potier	IFREMER	France	mikaela.potier@agrocampus-ouest.fr
Liesa Celie	The Flanders Research Institute for Agriculture, Fisheries and Food	Belgium	liesa.celie@ilvo.vlaanderen.be
Robert Thorpe	Centre for Environment, Fisheries and Aquaculture Science (CEFAS)	United Kingdom	robert.thorpe@cefas.gov.uk
Marie Savina Rolland	IFREMER	France	Marie.Savina.Rolland@ifremer.fr
Marianne Robert	IFREMER	France	Marianne.Robert@ifremer.fr
Roweena Patel	University of Reading	United Kingdom	r.patel@pgr.reading.ac.uk
Thomas Del Santo O'Neill	Queen Mary, University of London	United Kingdom	delsantooneillthomas@gmail.com
Maria Ching Villanueva	IFREMER	France	Ching.villanueva@ifremer.fr
Jamie Lentin	Shuttle Thread Limited	United Kingdom	lentinj@shuttlethread.com

Annex 2: WGSAM resolution

The **Working Group on Multispecies Assessment Methods** (WGSAM), chaired by Michael Spence, UK, will work on ToRs and generate deliverables as listed in the Table below.

	Meeting dates	Venue	Reporting details	Comments (change in Chair, etc.)
Year 2022	10-14 October	Woods Hole, USA	Reports on keyrun reviews to be provided after each review is complete	
Year 2023	9-13 October	Edinburgh, UK	Reports on any keyrun reviews that are completed	
Year 2024	7-11 October	Belfast, Northern Ireland, UK	Final report by 15 November to SCICOM	Outgoing co-chair: Valerio Bartolino, Sweden

ToR descriptors

ToR	Description	Background	Science Plan Codes	Duration	Expected Deliverables
a	Regional updates: Review further progress and deliver key updates on multispecies modelling and ecosystem data analysis contributing to modeling throughout the ICES region	This ToR acts to increase the speed of communication of new results across the ICES area	5.1; 5.2; 6.1	3 years	Report on further progress and key updates. Review and collaborate with appropriate EGs to revise sections on "species interactions" in the Fisheries Overviews
b	Key-runs: Parametrisation of multispecies and ecosystem key-run models for different ICES regions. This includes standard update (limited to inclusion of recent data), extensive update (incl. new data and processes), and new key-runs.	Key-runs are models checked against high quality criteria, which are developed to contribute to a variety of operational objectives as part of the ICES advice, i.e. provide information on natural mortality for inclusion in single species assessments, estimates of multispecies reference points, large operating ecosystem models for MSE, etc.	5.1; 5.2; 6.1	3 years	Report on output of multispecies models including stock biomass and numbers and natural mortalities for use by single species assessment groups and external users.
c	Skill assessment: Establish and apply methods to assess the skill of multispecies models intended for operational advice	This work is aimed at assessing the performance of models intended for strategic or tactical management advice.	5.1; 6.1; 6.3	3 years	Report on technical requirements for cross-models standardisation and comparison. Manuscript(s) on skill assessment of wide

		Evaluation will require work towards standardisation for cross-model comparison. This ToR will also deal with evaluation of methods for model calibration and data weighting in the context of multispecies modelling.			array of multispecies models based on a large simulation study.
d	Multi-model advice: Evaluate methods for generating advice by comparing and/or combining multiple models	This work is aimed at addressing structural uncertainty in advice arising from multiple models	5.1; 6.1; 6.3	3 years	Report on methods for comparing models and for constructing model ensembles. Report on case examples from both simulation testing and real studies
e	MSE: Evaluate methods and applications for multispecies and ecosystem advice, including evaluation of management procedures and estimation of biological reference points under the uncertainties of climate change.	This ToR looks for multispecies and ecosystem approaches to understand the resistance and resilience of ecosystems to a warming environment and to perturbations related to the effects of climate change. Through the use of simulations, alternative management strategies and exploitation regimes can be evaluated for robustness to uncertainties related to climate change.	2.5; 5.2; 6.1	3 years	Review methods to evaluate populations and ecosystem resilience. Review of methods for management strategy evaluation which incorporate the effects and uncertainties of climate change

Summary of the Work Plan

Year 1	All ToRs, update keyrun Baltic Sea (coupled with data preparation workshop for the Baltic Sea benchmark), keyrun Georges Bank multi-model (dedicated workshop)
Year 2	All ToRs
Year 3	All ToRs

Supporting information

Priority	The current activities of this Group will lead ICES into issues related to the ecosystem effects of fisheries under multiple sources of uncertainties incl. climate change. The activities will provide information (e.g., natural mortality estimates, performance of indicators, multispecies reference points) and tools (e.g., multi-model ensembles, keyrun models) valuable for the implementation of an integrated advice and the application of a precautionary approach in several North Atlantic ecosystems. Consequently, these activities are considered to have a high priority.
Resource requirements	The research programmes which provide the main input to this group are already underway, and resources are already committed. The additional resource required to undertake additional activities in the framework of this group is negligible.
Participants	The Group is normally attended by some 20–25 members and guests. Expertise in ecosystem dynamics, trophic interactions, modelling and fish stock assessment from across the whole ICES region.
Secretariat facilities	Standard EG support.
Financial	No financial implications.
Linkages to ACOM and groups under ACOM	ACOM, assessment Expert Groups.
Linkages to other committees or groups	WGMIXFISH, WGDIM, WGBIFS, IBTSWG, WGECO, all IEASG groups, WKCLIMAD.
Linkages to other organizations	None

Annex 3: Summary of presentations for ToR A

Robert Thorpe gave a presentation on the sensitivity of the North Sea StrathE2E2 model to the representation of zooplankton, and highlighted major impacts on biomass states, recovery timescales, and fisheries management, assuming a requirement for fisheries to achieve Good Environmental Status. The findings are of importance because zooplankton are a key intermediary between primary productivity and fish or fisheries but are often poorly represented in models.

The experimental design is summarised in Figure 1. 15 000 simulations cover 5 different rates of metabolism for omnivorous zooplankton (OZ), 3 different levels of warming (2003-13 reference climate, uniform 2C, 4C warmings), and 1000 different fishing scenarios for 3 fleets (demersal, pelagic, other gears). Each scenario is run for 200 years, 100 with constant fishing, followed by 100 years without fishing to assess recovery timescales.

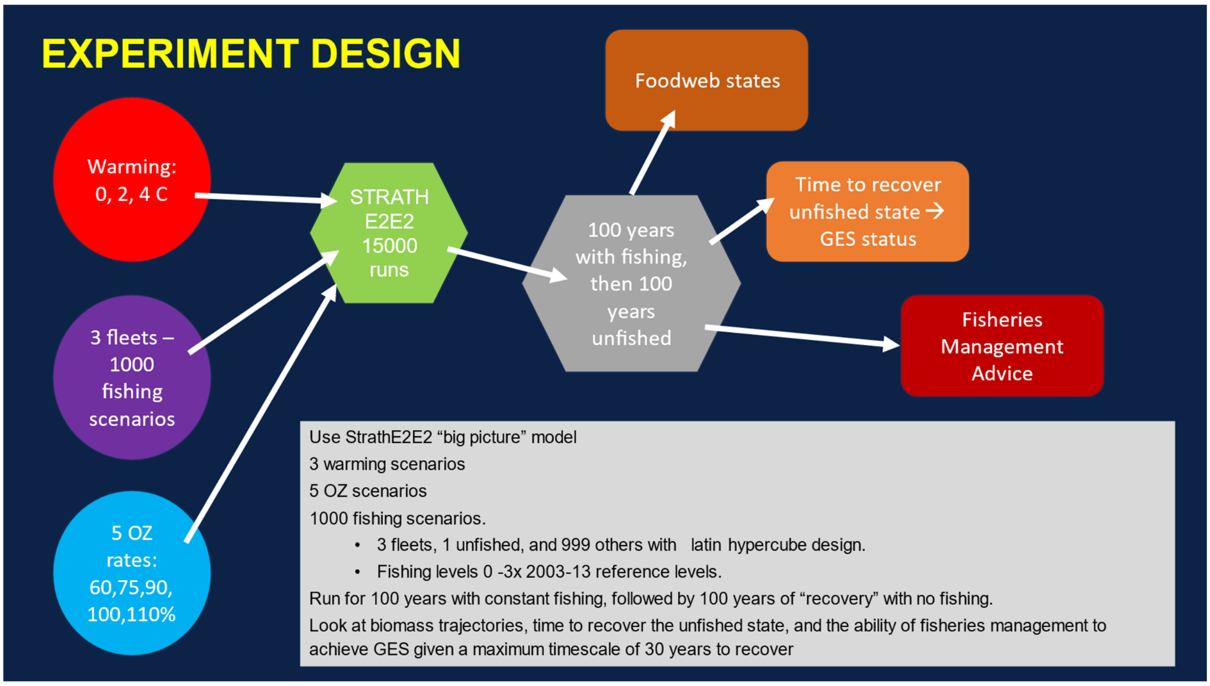


Figure 1. Schematic of experiment design.

We find that foodweb states are strongly influenced by both warming and OZ metabolism alongside fishing, but that OZ impacts exceed those of 4C warming (Figure 2).

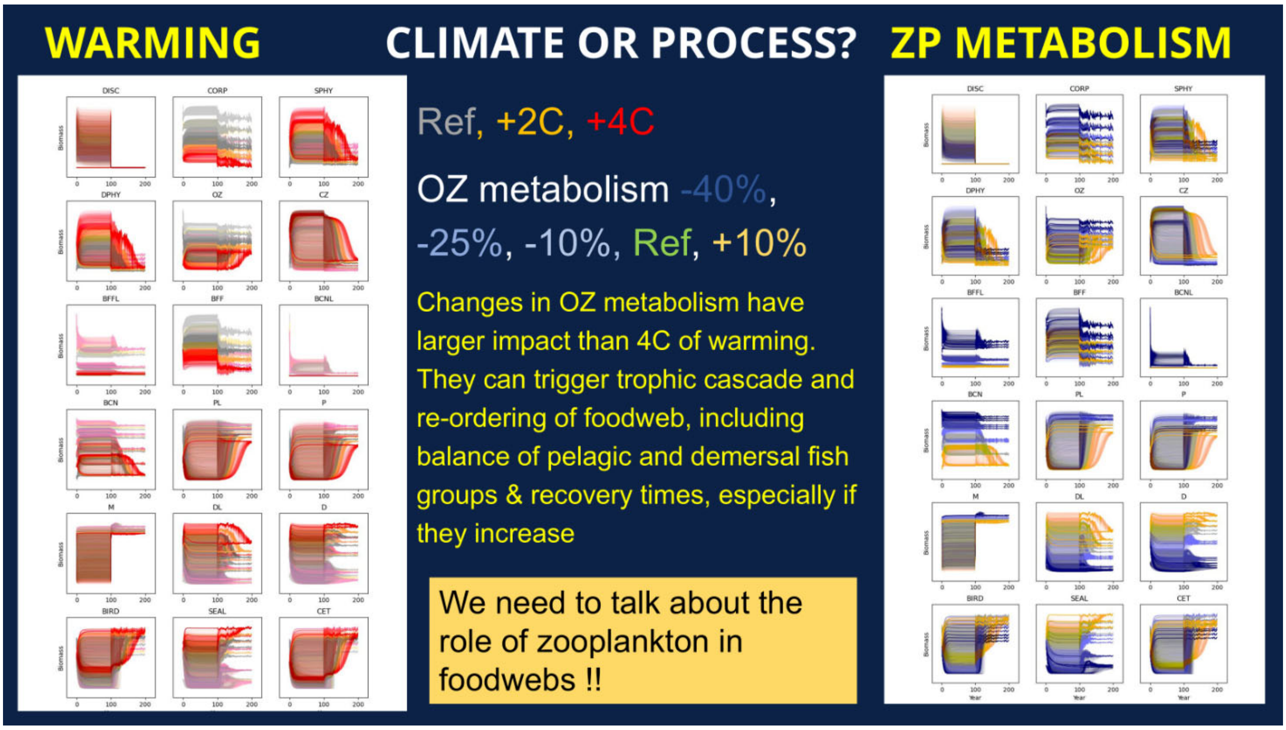


Figure 2. Relative impacts of warming and OZ metabolism.

OZ metabolism also strongly impacts recovery timescales. If OZ metabolism is high, OZ is less robust, and a trophic cascade that favours demersal over pelagic fish can be more easily induced, slowing recovery from fishing for pelagic, benthic, and OZ groups. If on the other hand, OZ metabolism is reduced, OZ is more robust, and the trophic cascade is less easily triggered, resulting in faster recovery times for mid-foodweb groups.

If we manage fisheries with the aim of achieving Good Environmental Status, this sensitivity would have strong impacts on fisheries management, in terms of whether fleet operations are more constrained for demersal or pelagic stocks (Figure 3).

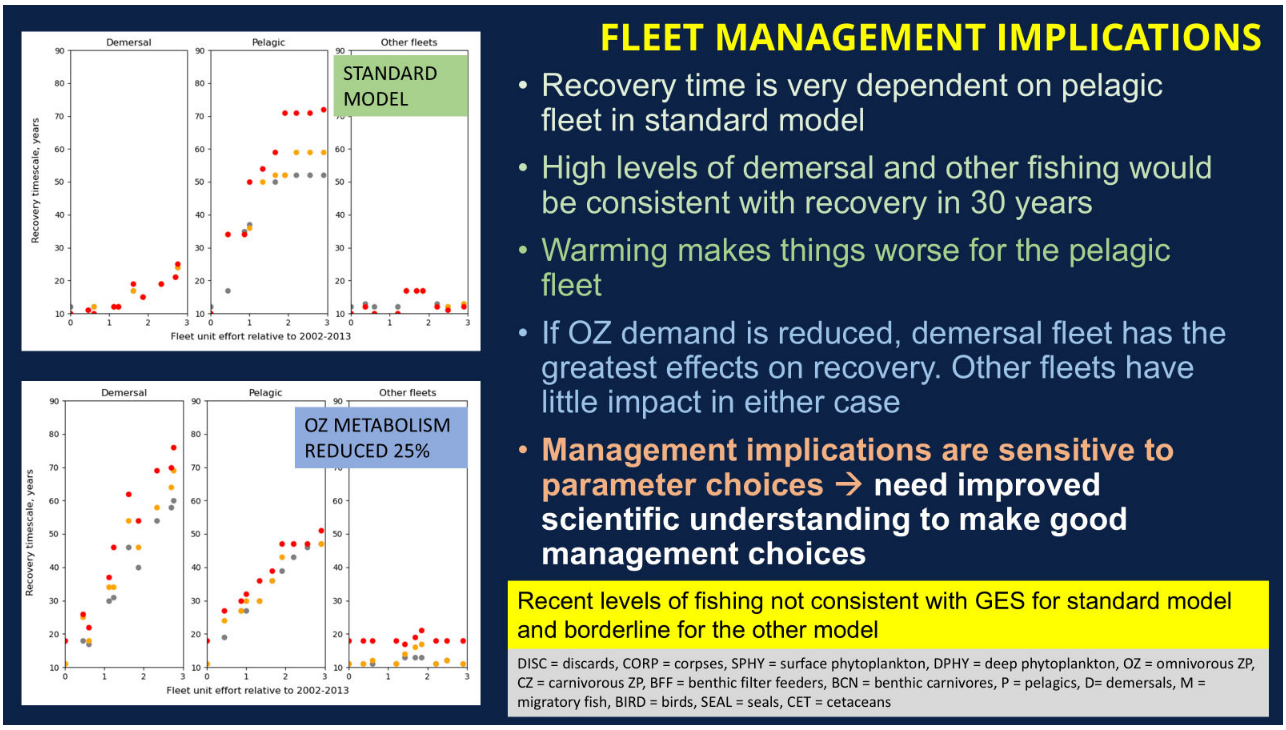


Figure 3. Fleet management implications of reduced OZ metabolism.

There are several caveats to the study. We only use one model, and this model has only a modest number of functional groups, and no size-structuring, or species, so the sensitivity could be a model artifact rather than a real property of the system. We also assume that a pattern of constant warming is reasonable, consistent with the CMIP ensemble. Subject to these caveats, we conclude that zooplankton can influence foodweb states, rates of recovery, and fisheries management targets, and that we need to improve our biological understanding of zooplankton and perhaps couple lower and higher trophic levels more in the future.

Including Seal Predation in Size-Structure Multispecies Models

Michael Thomson

Centre for Environment, Fisheries and Aquaculture Science (Cefas)

For ecosystems based fisheries management (EBFM) we need to model all interactions in an ecosystem rather than consider single species in isolation in order to more accurately determine how changing conditions/management affect all species and the relationship between them. For fisheries a comprehensive EBFM approach requires consideration of top level predators (marine mammals) together with fish species in multispecies ecosystem models used for fisheries.

We consider the situation of grey seals in the North Sea. There is evidence that grey seal numbers in the North Sea are increasing (Thomas *et al.*, 2019). The impact of this increase on fish species is uncertain and inclusion of grey seals in a North Seas multispecies fisheries model will allow quantification of how they are affecting the ecosystem.

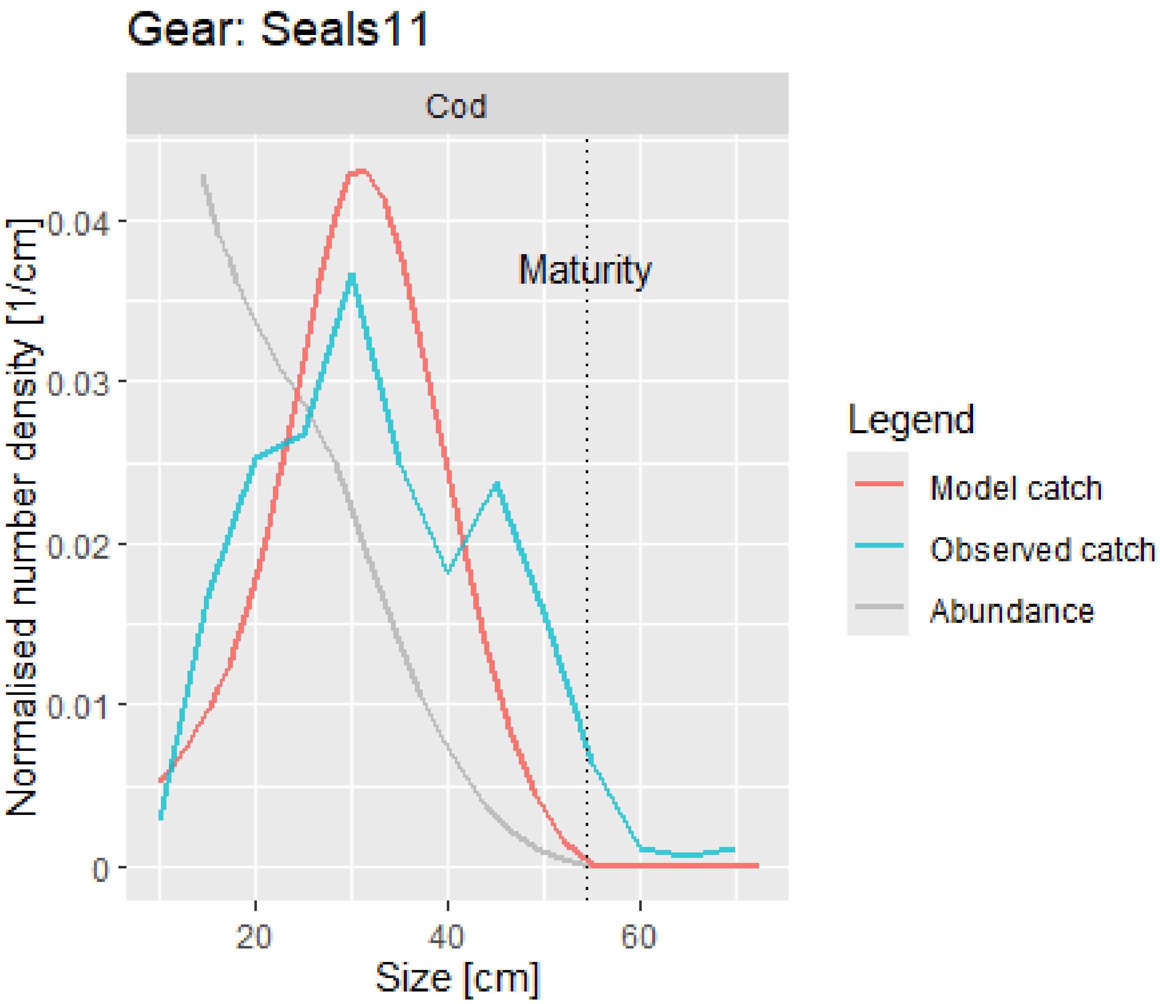
We are adding grey seals to mizer a dynamic multispecies size-spectrum model. (Scott *et al.*, 2014, Andersen *et al.*, 2016). Individual fish in the model are defined entirely by weight and species – tracked as they grow and progress through trophic levels. Mizer defaults are setup for fish species, it needs modification for marine mammals. Grey seals are added to an existing mizer model for the North Sea with 12 species included (sprat, sandeel, dab, herring, sole, haddock, plaice, gurnard, Norway pout, whiting, cod, saithe).

Thomas *et al.* (2019) have modelled seal population using a Bayesian state-space model fitted primarily to seal pup counts in the period from 1985-2010 using a particle filter. Fitted parameters in the model include adult survival, pup survival, fecundity, carrying capacity, sex ratio and shape of a density dependent survival function. Seal diet composition data in both species and length is obtained from Hammond & Wilson (2016) and Wilson & Hammond (2019).

As a first attempt a one-way feeding of the seal population model into mizer is made by including seals a "fishing fleet". A selectivity function and selectivity and catchability functions are fitted for each species in the model. To account for the effect of changing population effort is set as $F_{\{ij\}} = \gamma p_j N_i$ where $F_{\{ij\}}$is the seal "effort" on species j in year i, p_j is the proportion of seal diet that belongs to species j and N_j is the seal population in year i. The parameter γ determines how predation scales with seal population. All parameters are fitted by calibrating the model to catch estimates and the seal diet data. Calibration work is ongoing to improve model fit.

Alternative approaches to including seals are to either directly include as a species in mizer for which the default functions in mizer would need to be modified to account for the biology of marine mammals, or to introduce two-way coupling between the seal population model and mizer by keeping seals as a fishing fleet input to mizer and making the parameters of the seal population models functions of outputs of mizer (e.g. biomass of prey species) to introduce feedback between the two models.

Other mizer work being undertaken at Cefas are to add seabass (and possibly a recreational fishing fleet) to the North Sea mizer model, and end-to-end ecosystem modelling coupling the ERSEM model for lower trophic levels with mizer.



References

K.H. Andersen, N.S. Jacobsen and K.D. Farnsworth (2016). The theoretical foundations for size spectrum models of fish communities. Canadian Journal of Fisheries and Aquatic Science 73(4): 575-588.

P.S. Hammond and L.J. Wilson (2016). Grey Seal Diet Composition and Prey Consumption. Scottish Marine and Freshwater Science Vol 7 No 20, 47pp.

F. Scott, J.L. Blanchard and K.H. Andersen (2014) Mizer: an R package for multispecies, trait-based and community size spectrum ecological modelling. Methods in Ecology and Evolution 5(10): 1121-1125

L. Thomas, D.J.F. Russell, C.D. Duck, C.D. Morris, M. Lonergan, F. Empacher, D. Thompson and J. Harwood (2019). Modelling the population size and dynamics of the British grey seal. Aquatic Conservation: Marine and Freshwater Ecosystems 29(S1): 6-23.

L.J. Wilson and P.S. Hammond (2019). The diet of harbour and grey seals around Britain: Examining the role of prey as a potential cause of harbour seal declines. Aquatic Conservation: Marine and Freshwater Ecosystems 29(S1): 71-85

Multispecies models: US update

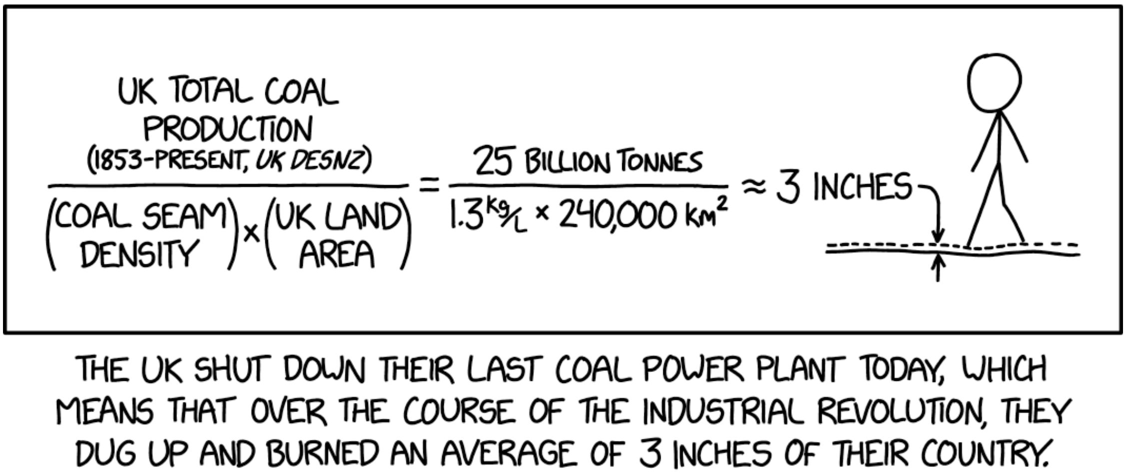
WGSAM ToR a, 7 October 2024

Sarah Gaichas, Alberto Rovellini, Andy Beet, Joe Caracappa, Gavin Fay (UMass Dartmouth), Robert Gamble, Max Grezlik (UMass Dartmouth), Isaac Kaplan, Emily Liljestrand, Sean Lucey (RWE), Maria Cristina Perez (UMass Dartmouth), James Thorson, Sarah Weisberg (Stony Brook U and NOAA), Robert Wildermuth

Updates

Since WGSAM 2023

- Atlantis climate and port integration, Northeast US (NEUS)
- Atlantis ecosystem yield (Gulf of Alaska)
- Rpath
    - Ascendancy and resilience
    - Portfolio management
- EcoState
- Single species applications
    - Stock assessment
    - Risk assessment

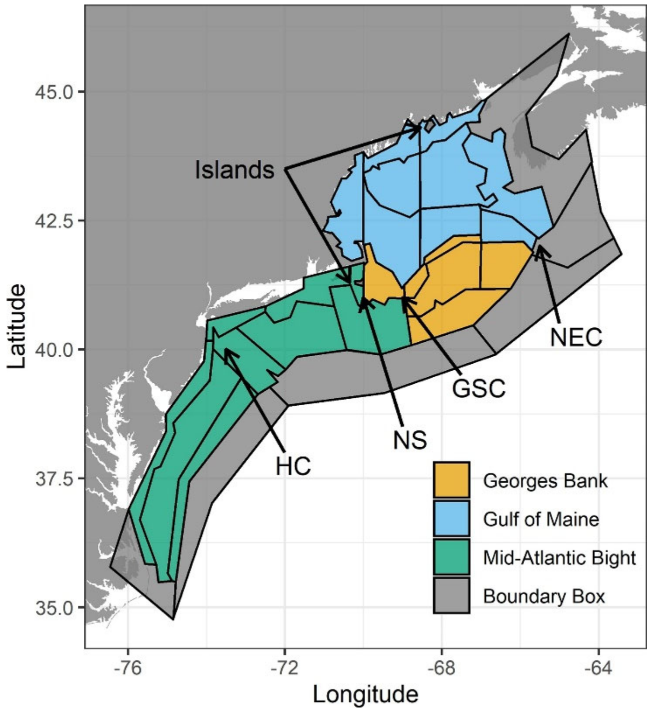


xkcd comic 2992 titled UK coal

.right[.contrib[One week ago today; https://xkcd.com/2992/]]

More Diverse uses of end to end and food web models

Atlantis ↓ and Rpath →



:img NEUSmap, 88%

Atlantis NEUSv2:
Updates for climate scenario testing

Joe Carracappa, Andy Beet, Robert Gamble

- Sensitivity to fishing scenarios (complete)
- Integrating spatial fleets and ports of origin
- Thermal thresholds project to integrate species temperature ranges
- Climate projections using MOM-6 planned
- Testing ecosystem overfishing indicators for the Northeast US

.footnote[https://github.com/NOAA-EDAB/neus-atlantis]

Albi Rovellini's ongoing work on evaluating GOA optimum yield with Atlantis. This is part of a study on multispecies fishing simulations under different climate regimes and fishing configurations.

- Gulf of Alaska groundfish is managed with a multispecies ecosystem cap on annual catch allocations
- Optimum Yield cap = 800,000 mt each year, sum of single-species ABC must be lower
- Similar system in the Bering Sea, but higher cap (2M t)
- Since 1987, the GOA OY cap has never constrained catch allocations (unlike the Bering Sea cap), suggesting that the OY cap may be too high for the productivity of GOA stocks
- Given its history and projected climate change in the GOA, should we expect the OY cap to constrain GOA catches in the future?
- Multispecies fishing simulations with Atlantis, 4 climate-fishing combinations:

- Fishing: (1) all stocks fished at equal multipliers of FMSY; (2) same as (1) except arrowtooth flounder is only lightly exploited
- Climate: from ROMS, (1) pre-heatwave 1999 cool conditions; (2) high CO2 emissions, end of century scenario

Results: Underexploitation of key groundfish predator Arrowtooth flounder (right-hand column) leads to lower global yield, because arrowtooth predates on groundfish (mostly walleye pollock). Arrowtooth flounder is currently lightly exploited because it has limited commercial value Warmer climate (bottom row) leads to lower global yield, largely because of near collapse of Pacific cod under warm conditions mediated by recruitment failure

TAKE HOME MESSAGE: The OY cap in the GOA is unlikely to constrain fishery allocations in the future

Rpath: Relative ascendency for three adjacent food webs

Sarah Weisberg PhD thesis

Highlight key groups within food webs and identify regime shifts

- Ecological network analysis shows different potential resilience across MAB, GB, GOM, and regimes in Gulf of Maine food web efficiency/resilience.

- Benthic vs pelagic groups across the three systems

- Highly efficient food webs have lower resilience due to fewer trophic pathways decreasing redundancy.

- The Gulf of Maine had low resilience in the 2000s, corresponding to poor fish condition

.footnote[p.s. Shiny GOM used in IEA course]

Rpath: Supporting Portfolio Theory Management of Georges Bank Fisheries

Max Grezlik PhD Thesis

Efficiency frontiers quantify revenue and financial risk taken to achieve a given revenue.

Red and blue lines compare an expanded groundfish complex to single species management of the same species.

The point is the observed revenue and risk of forgone revenue for a given year.

Preliminary results suggest the current, 13-species complex allows for some benefits beyond what single species management would achieve. The distance between observed and EBFM frontier suggests there is added economic benefit to expanding the current complex.

Goal: demonstrate the utility of diversification of fishing portfolios in New England

- NEFMC efforts toward climate-resilient fisheries.
- Worked example: expanding Northeast Groundfish Complex.
- Fishing portfolio diversification empowers fishers to find solutions to choke stocks which have threatened the sustainability and profitability of an historic fishery.

Efficiency frontiers quantify revenue and financial risk taken to achieve a given revenue. Red and blue lines compare an expanded groundfish complex to single species management of the same species. The point is the observed revenue and risk of forgone revenue for a given year. Preliminary results suggest the current, 13-species complex allows for some benefits beyond what single species management would achieve. The distance between observed and EBFM frontier suggests there is added economic benefit to expanding the current complex.

Georges Bank Keyrun Review and Work in progress

- Challenge of place based approach for stocks with substantial dynamics outside Georges Bank: "In that case, expanding the models outside the boundaries of the EPU, and/or explicitly accounting for the input/output of fish and energy across the boundaries will likely be needed"
- Dedicated R packages for data positively reviewed
- Standardize diet interactions and better quantify other food in estimation models using Rpath
- Do model self-tests
- Model specific structural and sensitivity recommendations

In progress

- Self tests (4 species Hydra in progress by Cristina)
- Model specific recommendations
    - Fleet changes done, 3 fleet being implemented by Emily L
    - Feeding parameters done
- Testing stalled but hope to resume (ToR c)
- Work continues on input datasets (landings and discards)

EcoState: new state-space dynamic food web model

Jim Thorson's Paper in review: https://ecoevorxiv.org/repository/view/7476/

Simulation test

Single species applications

Time varying natural mortality for Atlantic mackerel

Ecological drivers of recruitment for Atlantic herring

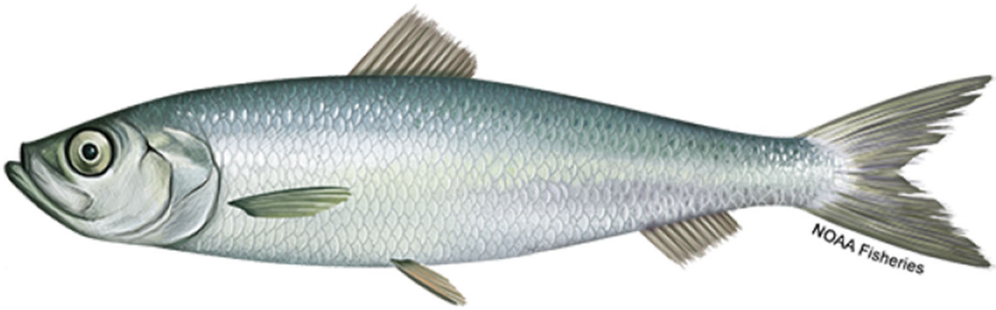
Laurel Smith *et al.*

Does food drive recruitment of Atlantic herring?

Create zooplankton indices to evaluate changes in food for Atlantic herring larvae, juveniles, and adults over time and in space in the Northeast US continental shelf ecosystem.

Two applications:

1. Addressing uncertainty in the stock assessment for Atlantic herring (*Clupea harengus*)
2. Describing zooplankton species and group trends for integrated ecosystem assessment



Atlantic herring illustration, credit NOAA Fisheries

.footnote[https://www.naturepl.com/stock-photo-copepods-calanus-finmarchicus-aggregated-at-the-surface-reverse-diel-nature-image01407135.html]

Which indicators are relevant for recruitment?

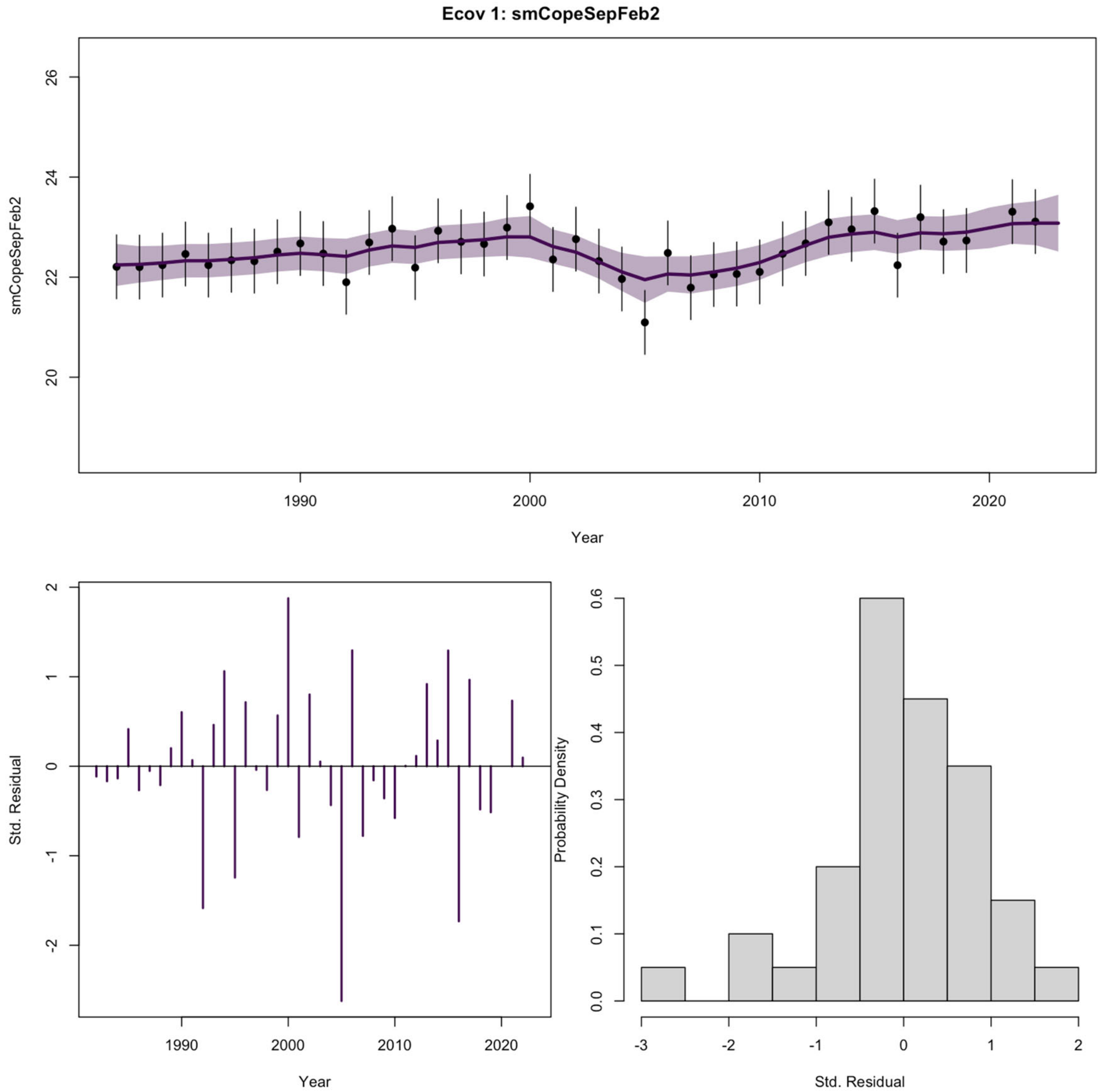
Which indicators are potential covariates for recruitment?

Boosted regression tree (Molina 2024) investigated relationships between environmental indicators and Atlantic herring recruitment estimated in the assessment.

Larval and juvenile food (zooplankton), egg predation, and temperature always highest influence

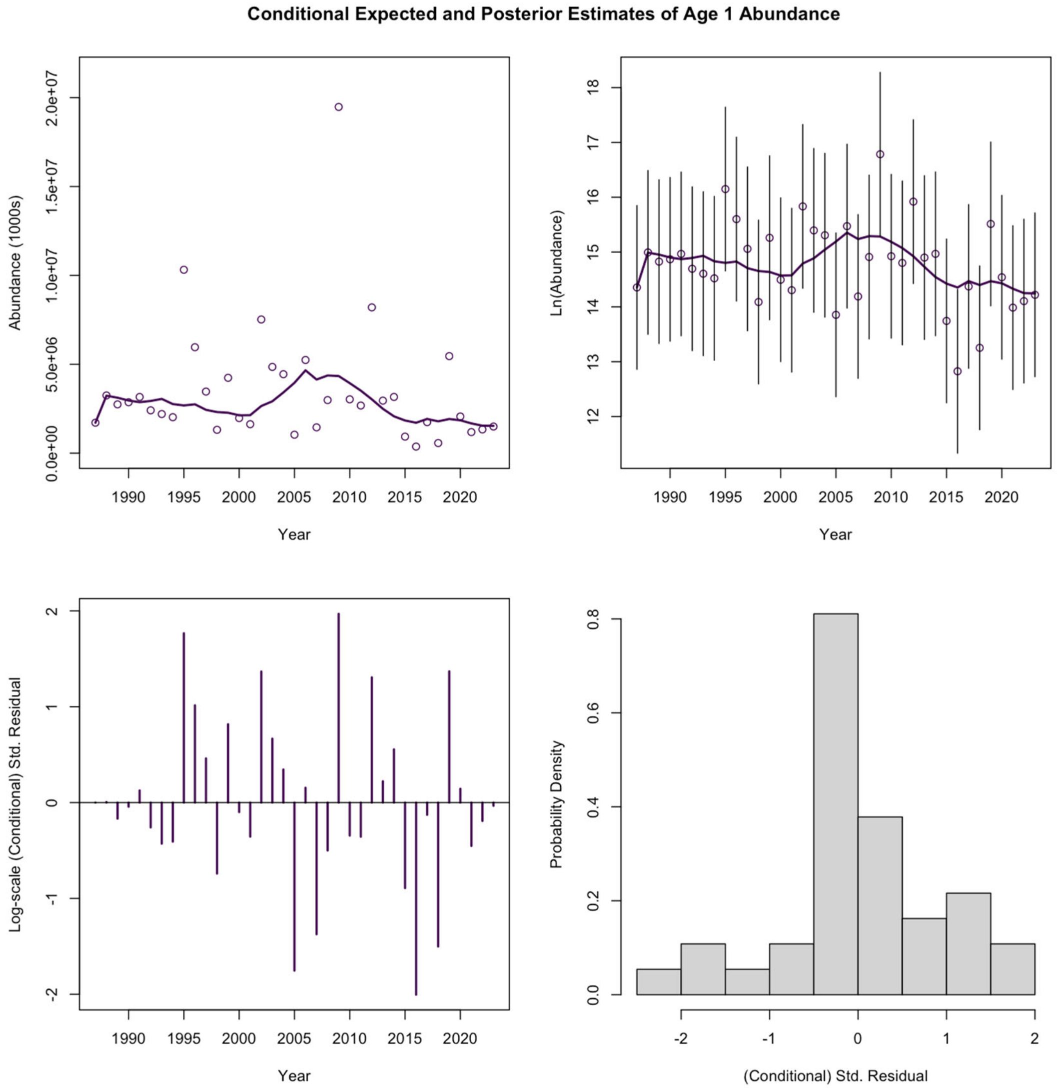
How to include in the stock assessment?

Zooplankton example



]

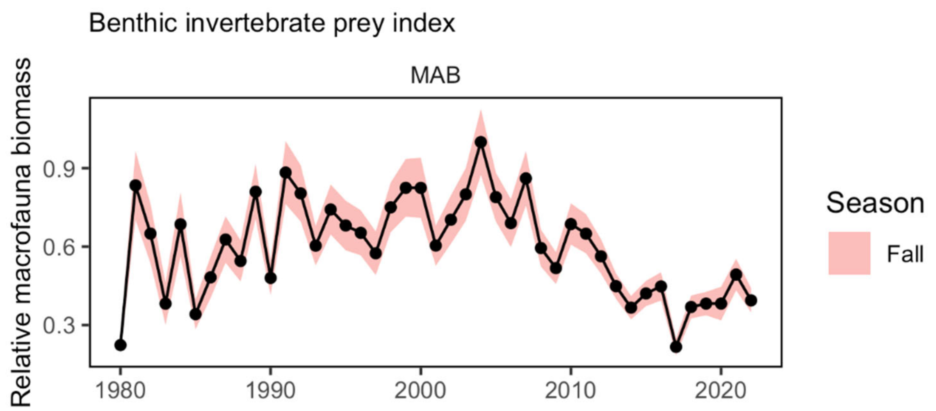
The inclusion of the zooplankton index improved model fit IN THE WRONG DIRECTION.

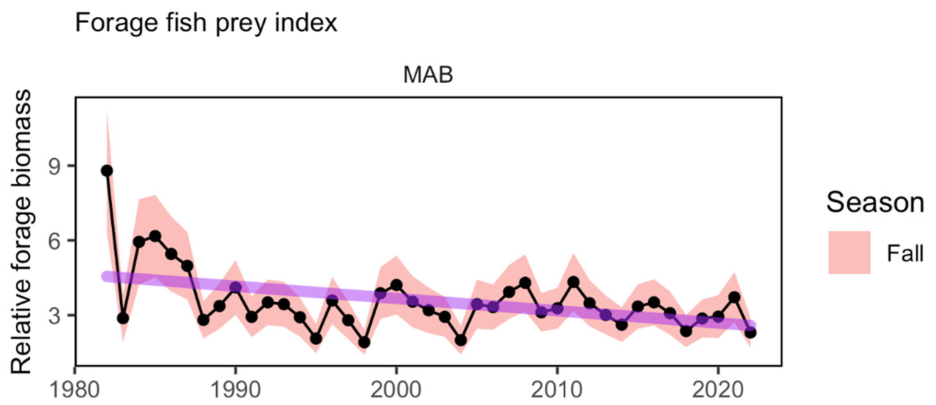


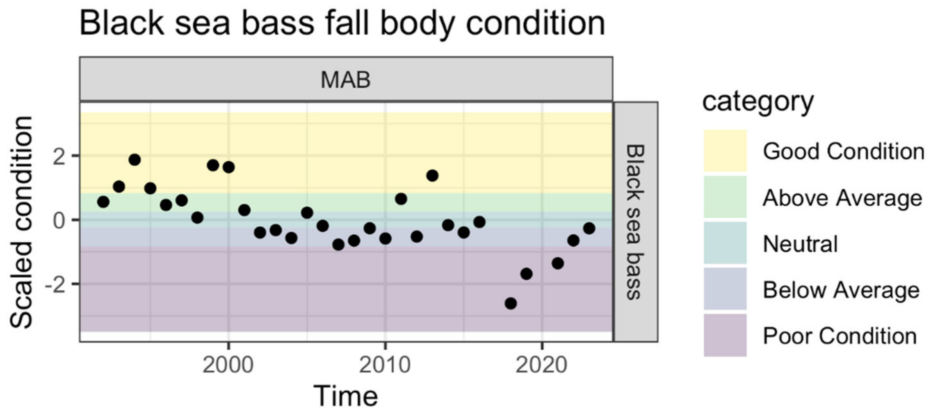
Haddock egg predation and thermal habitat better

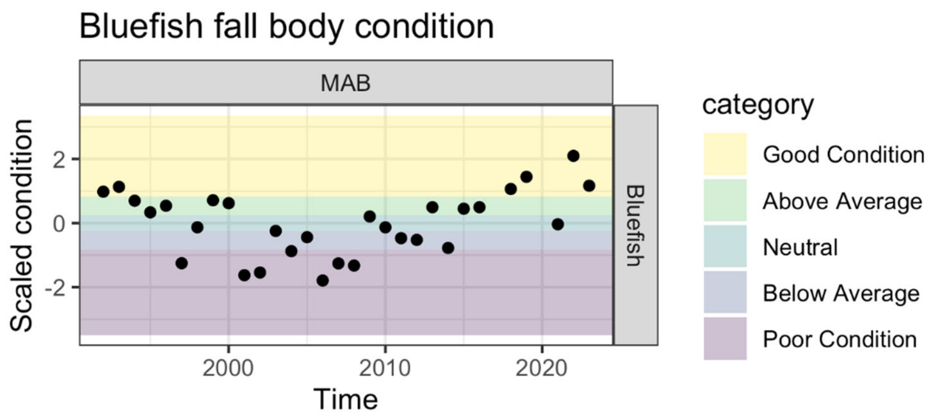
Risk assessment applications for fishery managers

Example: Evaluate risks posed by prey availability to achieving OY for Council managed species









Council and Advisory Panel members recommended new elements addressing human dimensions (recreational access equity), new elements addressing cross-sectoral impacts (offshore wind impacts on biology and ecosystem as well as fishery access and scientific sampling), and transitions from static ecosystem indicators to time series indicators (prey availability, predation pressure, and fishing community vulnerability). New ecosystem science was required to support these requests. The process included development of new indicators of prey availability based on spatio-temporal modeling using ecological datasets (stomach contents, zooplankton), and new spatial analyses of habitat, revenue, and surveys relative to wind energy development areas. Development of potential risk criteria is ongoing; thresholds between low, moderate, and high risk that are essential to operational use are developed collaboratively with Council and Advisory Panel members.

The slide shows a higher risk example (black sea bass, low recent condition correlated with recently declining prey) and a lower risk example (bluefish, despite a long term decline in forage fish prey. recent condition has been good)

Review of available marine ecosystem models (MEMs) applied to French sea waters (FORESEA project)

Morgane Travers-Trolet, Raphaël Girardin, Sigrid Lehuta

Funded by Ifremer, the FORESEA project (*French seafOod pRoduction Scenarios in 2050*) aims at integrating available knowledge across disciplines and scales and filling knowledge gaps in order to build plausible foresight scenarios of French commercial fisheries under global change by 2050 and predict possible pathways of domestic marine ecosystems based on those scenarios. After a first step of development of the foresight scenarios, the project aims a simulating them using a suite of marine ecosystem models applied to the French waters. A call to modellers has therefore been launched in the form of an online survey with very little constraints on the type of models, except that they must be able to make projections including the impacts of climate change and fishing in 2050. The questionnaire (https://forms.ifremer.fr/foresea/models-review/) is structured in 5 pages to inform on general characteristics of the model, details about components and processes, how fishing and climate are modelled, calibration and output and capabilities of the model to be used within ForeSea.

It was the opportunity for a synthesis of modelling capacities within French waters, and evidence a large spatial coverage by a diversity of models dominated by ecosystem models (9), multispecies size-based models and fleet and management models (35 entries). In general, the review evidenced a poor representation of zooplankton diversity and on the other hand of the line, heterogeneous ways to account for fishing fleets. Fifty percent of the models had climate change scenarios ready to exploit mainly for those covering the low trophic levels. On the contrary, very few fleet/management scenarios were available at the horizon 2050, mostly in a stylized manner. The exercice evidenced the challenge of modelling the socio-ecosystem and the need to combine models to quantitatively cover the richness of the visions of the future provided by the foresight scenarios.

2022

FishStomachs, an R-Package for compilation of stomach contents data for estimation of population diet and food ration.

Morten Vinther, DTU Aqua, Denmark

Compilation of data from individual stomachs to population level is not trivial and the method will depend highly on the question asked. Stomach data are often collected to get information on the average diet or food ration of a given species within a given area to inform multispecies assessment models (e.g. SMS or Gadget). Such models require data on diet and biomass eaten (food ration) by predator species and size classes. The average “population” diet or food ration should basically be calculated from a stratified mean of the individual stomach content samples, weighted by strata density of the predator and the area of the strata. This seems simple, but incomplete and patchy sampling makes it often necessary to use a series of *ad hoc* solutions.

In general, a series of data compilation methods could be applied:

1. Read and check data from agreed exchange format;
2. Bias correct to take into account variable evacuation rate;
3. Assign size classes for predators and preys;
4. Bias correct to take into account regurgitated stomachs within sample units;
5. Aggregate stomachs contents within sample_id and size classes.

6. Allocate unidentified or partly identified prey items to fully identified items;
7. Calculate population diet and food ration from a weighted average.

The sequence of these steps, of which some of the steps might be repeated, will depend on the individual sampling design and the quality of the analysed data.

The FishStomachs R-package defines data structures suitable for stomach data, and provides the necessary methods to compile observed stomach data into population diet and biomass eaten, used for multispecies models. The methods applied for a set of observations are stored within the data output to document the compilation steps taken.

"FishStomachs" is available from https://github.com/MortenVinther/FishStomachs

WGSAM 2022 US ToR a report

Sarah Gaichas, Grant Adams, Kerim Aydin, Brandon Beltz, Pierre-Yves Hernvann Kirstin Holsman, Isaac Kaplan, Sean Lucey, Janet Nye, Jameal Samhouri, Andy Whitehouse, Sarah Weisberg

2022-10-17

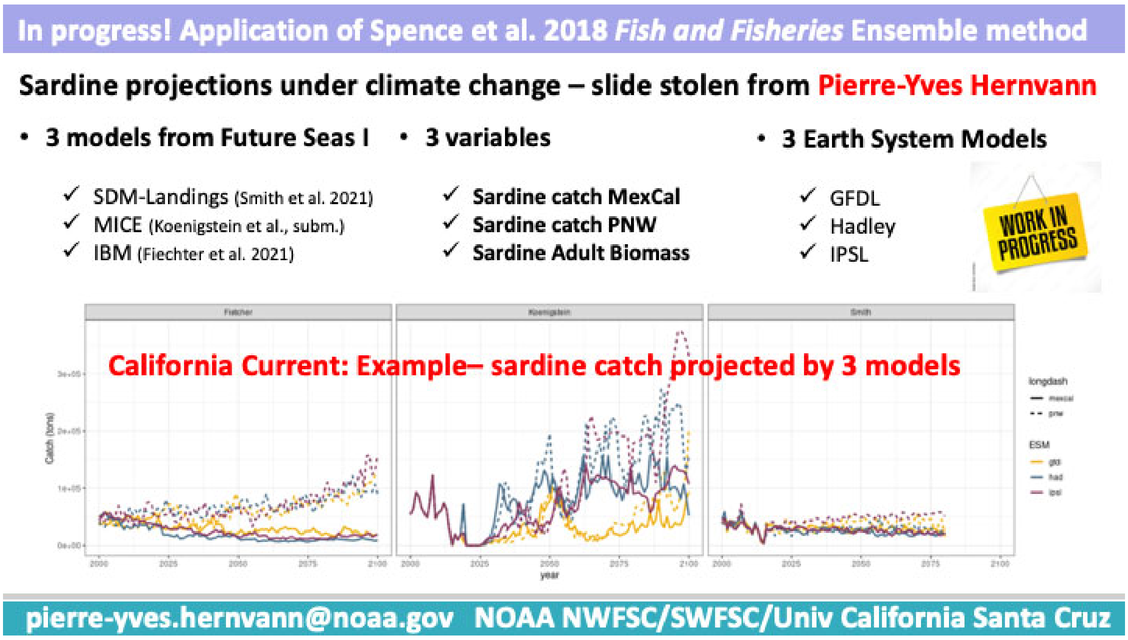
ToR a US Report

Sarah Gaichas presented a summary of ongoing multispecies and ecosystem modeling work from around the US.

Two new publications from Alaska highlighted new applications of multispecies (Adams *et al.* 2022) and ecosystem (Whitehouse *et al.* 2021) models. The multispecies model CEATTLE (Kirstin K. Holsman *et al.* 2016; K. K. Holsman *et al.* 2020) has been re-implemented in Template Model Builder (TMB, Kristensen *et al.* (2016)) as the R package Rceattle, available at https://github.com/grantdadams/Rceattle. The application of Rceattle in the Gulf of Alaska was used to compare consumption estimates for key groundfish species within an ensemble of differently parameterized models. MSE work is also ongoing with the Rceattle model in the Gulf of Alaska. It is likely that the Bering Sea 3 species model presented alongside the walleye pollock assessment will now be deployed in RCeattle. The full food web model Rpath, available at https://github.com/NOAA-EDAB/Rpath was used with a suite of climate models to evaluate potential future food web impacts of climate change, and how these impacts would interact with Bering Sea fishery management (Whitehouse *et al.* 2021). More information is available at Alaska multispecies and ecosystem models.

One new publication from the US Great Lakes implemented a multispecies assessment for two predators, Chinook salmon and lake trout, with their common prey Alewife in Lake Ontario (Fitzpatrick *et al.* 2022). This model linked Alewife (prey) biomass to predator growth, which has been of interest to WGSAM for many years. Each species was modeled as with typical catch at age dynamics equations, as well as bioenergetics models for predators that linked alewife consumption to predator growth. The feedback between prey and predator growth was important to explain dynamics and to provide insight into the effectiveness of potential management responses (e.g., reduce stocking of one or more predator species to maintain prey biomass).

On the US West Coast (California Current ecosystem) the FutureSeas project is ongoing to link climate projections to food web responses, including key forage and predator species. More information is available at California Current Future Seas MSEs. Pierre-Yves Hernvann is in the process of combining results from several models into an ensemble using approaches presented at WGSAM previously (Spence *et al.* 2018).



Ensemble modeling in FutureSeas project

In the Northeast US, Georges Bank modeling efforts will be reported under ToRs b and c. There are also two new Rpath models for the Mid-Atlantic Bight and the Gulf of Maine. Mid-Atlantic Bight work by Brandon Beltz (MS thesis, 2022) included building the initial model, then applying forced migration for spiny dogfish and other sharks to examine the impacts of climate-driven changes in predator migration and timing: simulations included changing the amount leaving the system (Range shift), changing the timing of migration (Phenological shift), and a combined Range and Phenological shift. Sarah Weisberg (PhD thesis, ongoing) presented work based on the Gulf of Maine Rpath model at ICES WGCOMEDA in October 2022. Her work included building the initial model, code improvements to implement trophic network metrics related to resilience such as System Omnivory, Finn's Cycling, and Relative Efficiency and Redundancy. Ongoing work is using Ecosense to generate distributions of systems to classify systems into low and high efficiency, characterize biomass distributions in high and low efficiency systems, and to evaluate changes in low and high efficiency system states over time.

References

Adams, Grant D., Kirstin K. Holsman, Steven J. Barbeaux, Martin W. Dorn, James N. Ianelli, Ingrid Spies, Ian J. Stewart, and André E. Punt. 2022. "An Ensemble Approach to Understand Predation Mortality for Groundfish in the Gulf of Alaska." *Fisheries Research* 251 (July): 106303. https://doi.org/10.1016/j.fishres.2022.106303.

Fitzpatrick, Kimberly B., Brian C. Weidel, Michael J. Connerton, Jana R. Lantry, Jeremy P. Holden, Michael J. Yuille, Brian Lantry, *et al.* 2022. "Balancing Prey Availability and Predator Consumption: A Multispecies Stock Assessment for Lake Ontario." *Canadian Journal of Fisheries and Aquatic Sciences* 79 (9): 1529–45. https://doi.org/10.1139/cjfas-2021-0126.

Holsman, K. K., A. C. Haynie, A. B. Hollowed, J. C. P. Reum, K. Aydin, A. J. Hermann, W. Cheng, *et al.* 2020. "Ecosystem-Based Fisheries Management Forestalls Climate-Driven Collapse." *Nature Communications* 11 (1): 4579. https://doi.org/10.1038/s41467-020-18300-3.

Holsman, Kirstin K., James Ianelli, Kerim Aydin, André E. Punt, and Elizabeth A. Moffitt. 2016. "A Comparison of Fisheries Biological Reference Points Estimated from Temperature-Specific Multi-Species and Single-Species Climate-Enhanced Stock Assessment Models." *Deep Sea Research Part II: Topical Studies in Oceanography* 134 (December): 360–78. https://doi.org/10.1016/j.dsr2.2015.08.001.

Kristensen, Kasper, Anders Nielsen, Casper W. Berg, Hans Skaug, and Bradley M. Bell. 2016. “TMB: Automatic Differentiation and Laplace Approximation.” *Journal of Statistical Software* 70 (April): 1–21. https://doi.org/10.18637/jss.v070.i05.

Spence, Michael A., Julia L. Blanchard, Axel G. Rossberg, Michael R. Heath, Johanna J. Heymans, Steven Mackinson, Natalia Serpetti, Douglas C. Speirs, Robert B. Thorpe, and Paul G. Blackwell. 2018. “A General Framework for Combining Ecosystem Models.” *Fish and Fisheries* 19 (6): 1031–42. https://doi.org/10.1111/faf.12310.

Whitehouse, George A., Kerim Y. Aydin, Anne B. Hollowed, Kirstin K. Holsman, Wei Cheng, Amanda Faig, Alan C. Haynie, *et al.* 2021. “Bottom–Up Impacts of Forecasted Climate Change on the Eastern Bering Sea Food Web.” *Frontiers in Marine Science* 8. https://www.frontiersin.org/articles/10.3389/fmars.2021.624301

Impacts of Warming and Fishing on Functional Group Abundances and Timescales of Recovery from Fishing

Robert Thorpe gave a presentation on work carried out with Mike Heath’s r-package StrathE2E2 in the North Sea under Tor A. Two pieces of work were presented, firstly a published study on the effects of warming and fishing on North Sea functional groups, and secondly a recent exploratory study looking into the use of recovery timescales as means of determining whether Good Environmental Status (GES) is being achieved in the North Sea.

StrathE2E2 (Heath *et al*. 2020; 2021) is a MICE end-to-end model of the ecosystem, which uses nitrogen as its prime currency of exchange (like Atlantis). It has been designed for rigorous evaluation of broad-scale top down and bottom up effects, and has 18 broad functional groups and 5 resource pools. Representation of feeding, metabolism, reproduction, migration, advection, and mixing are included, and the model is driven by monthly chemical and physical boundary conditions from oceanographic models, so a seasonal response is included.

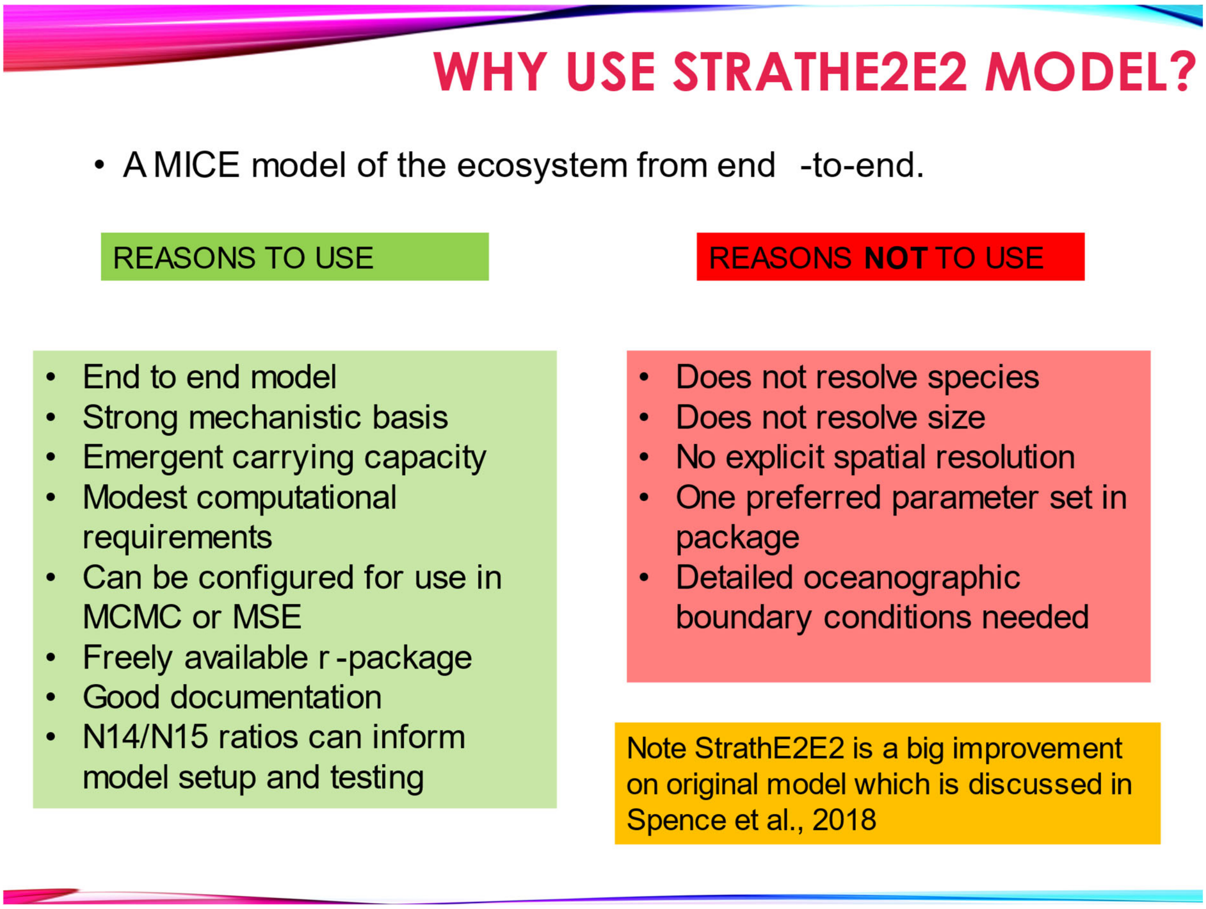


Figure 1: Key features of the distinctive StrathE2E model framework.

In these studies we used the model to look at broad-scale impacts of warming and fishing on the ecosystem form and function across the 18 functional groups, and applied this to the management question as to whether GES was being achieved.

In the first study (Thorpe *et al.*, 2022), we considered 16 fisheries scenarios, covering two main fleets demersal, and pelagic, and considering effort levels from zero up to twice the 2003-13 period (heavy fishing), alongside 3 warming scenarios, a reference climate (continuation of 2003-13) and warming of 2K or 4K. We found that warming increased productivity (in contrast with some studies – Thorpe *et al.*, 2022, but supported by others (van Leeuwen at al., 2021)) with little change in function, whilst fishing changed energy flows and function, but had little effect on overall size. In particular fishing of pelagics could trigger a cascade in which their numbers decreased relative to demersals, with warming increasing the likelihood of such a cascade. Because the effects of warming and fishing were so different, our study found that monitoring of broad biomass groups could help with untangling their effects and with determination of GES.

In Thorpe *et al.*, 2022 we suggested that reference states for GES should be defined relative to unfished (dynamic B0) and that this could be done either in distance of biomass ratios from unfished or in terms of time to recovery (Rossberg *et al.*, 2017). In the second part of this study we assessed the latter method by looking at the recovery trajectories of 1000 fishing scenarios for the 3 warming scenarios (reference, 2K, and 4K warming). Fishing scenarios involved independently varying the effort of 3 fleets (demersal, pelagic, and other) between zero and 3x the average intensity of the recent (2003-13) past. Scenarios were run for 100 years under the chosen constant fishing pressure, and 100 years of no fishing to allow the exploration of recovery rates.

We asked 5 questions relevant for the application of the Rossberg *et al.* (2017) method. 1) does a focus on recovery along permit present states so impacted they should be ruled out on other grounds, 2) to what extent does typical variability impact recovery timescales, 3) are results sensitive to the manner in which we define recovery, 4) and relatedly, how sensitive are results made about the "no fishing" baseline against which recovery is judged. Finally (5) we asked how sensitive the results were to the previously established trophic cascade (Thorpe *et al.*, 2022).

Our results confirmed that the suggested recovery timescale was reasonable – much longer than that would allow severe depletion to occur. We found that the recovery timescale was mostly set by the pelagic fleet in the standard model – demersal and other fleets could fish at high intensity whilst still permitting a rapid recovery. This was due to the pelagic trophic cascade; if this was reduced by changing parameter settings, recovery became more dependent on levels of demersal fishing. We found that the warmer scenarios were slower to recover. Natural variability and definition of recovery were found to have relatively minor impacts, but the impact of shifting baselines was substantial (Figure 2), with a 2K warming making it impossible to recover the unfished reference state, irrespective of management action.

Only one model was used in the study, so we have not been able to evaluate structural uncertainty (Spence *et al.*, 2018) whilst parameter uncertainty appears important. However despite these and other caveats (weaknesses in Figure 1), we suggest that the Rossberg method is potentially useful for detemination of GES, particularly when used alongside methods that evaluate the current state. If this method is used, it will have implications for model development and calibration, because timescales of model response become as important as state trajectories.

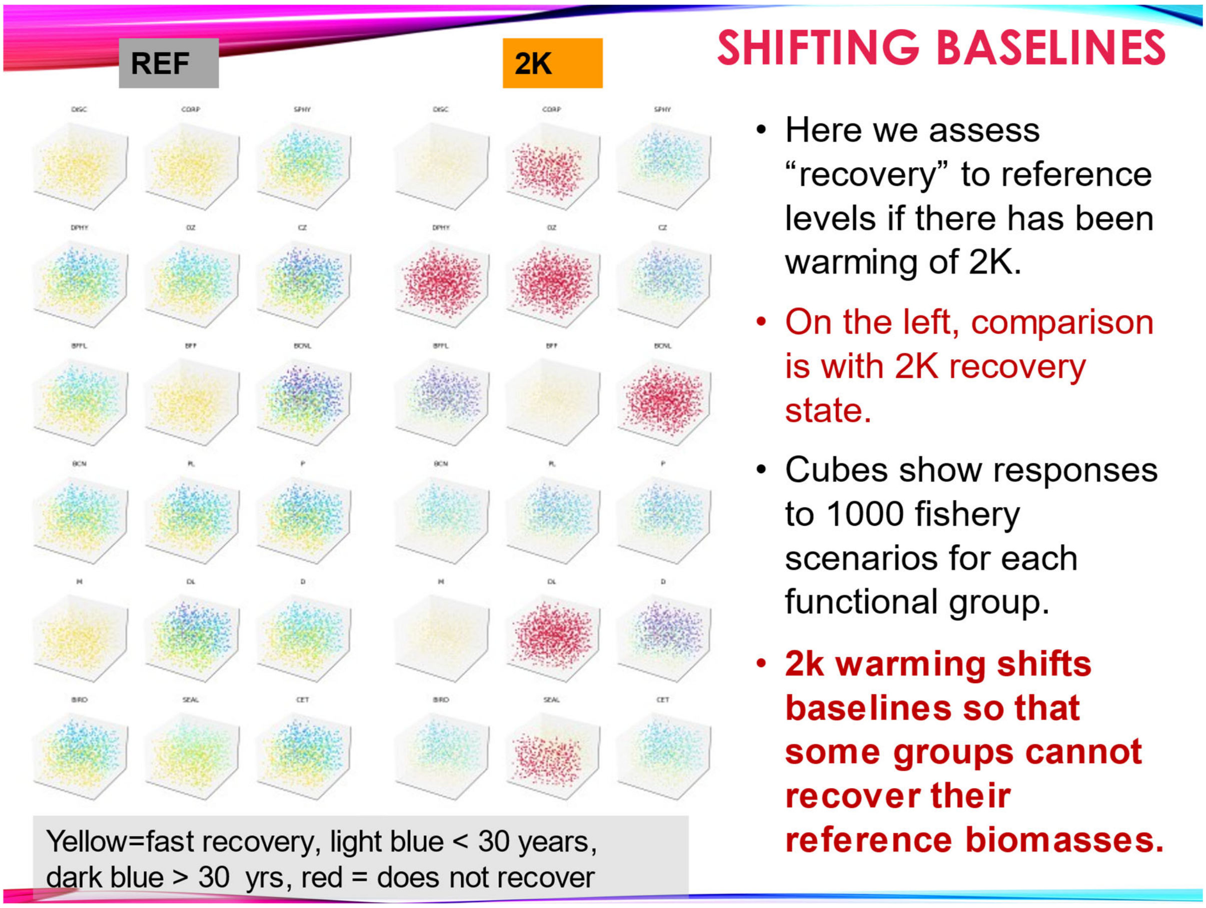


Figure 2: Impact of shifting baselines for a warming of 2K. On the left is the reference climate response, on the right is the response to 2K warming, when we assume the reference recovery state is still valid.

References

Heath, M. R., Speirs, D. C., McDonald, A., and Wilson, R. (2021). StrathE2E2 Version 3.3.0: Implementation for the North Sea. Available online at: https://marineresourcemodelling.gitlab.io/ (accessed January 31, 2022).

Heath, M. R., Speirs, D. C., Thurlbeck, I., and Wilson, R. (2020). StrathE2E2: an R package for modelling the dynamics of marine food webs and fisheries. Methods Ecol. Evol. 12, 280–287. doi: 10.1111/2041-210X.13510

Rossberg, A.G., Uusitalo, L., Berg, T., Zaiko, A., Chenuil, A, Uyarra, M.C., Borja, A., Lynam, C.P. 2017. Quantitative criteria for choosing targets and indicators for sustainable use of ecosystems, Ecological Indicators, 72, 215-224. http://dx.doi.org/10.1016/j.ecolind.2016.08.005

Spence, M. A., Blanchard, J. L., Rossberg, A. G., Heath, M. R., Heymans, J. J., Mackinson, S., *et al.* (2018). A general framework for comparing ecosystem models. Fish Fish. 19, 1031–1042. doi: 10.1111/faf.12310

Thorpe *et al.*, 2022. The response of North Sea functional groups to warming and changes in fishing, Front. Mar. Sci. 9, 841909

Assessment workflow tools & the MFDB package

Jamie Lentin (Shuttlethread, UK), Bjarki Þór Elvarsson (MFRI, Iceland), Will Butler (MFRI, Iceland)

https://presentations.shuttlethread.com/2022-10-10-assessment-workflow-tools-mfdb.html

Assessment models have increasing demands for data from disparate sources; historical commercial logbooks, open data APIs, and internal institutional databases, and aggregating this data has become a computational challenge in itself.

Using tools such as duckdb (https://duckdb.org) and dbplyr (https://dbplyr.tidyverse.org/) provide a very simple way to manage very large amounts of data, and allow the aggregation process to be scripted, and thus repeated and further developed easily. However, offer few mechanisms for managing more complicated structures than can be represented in flat tables.

MFDB (https://gadget-framework.github.io/mfdb/) is a data-handling package written in R, available on CRAN. It offers a "pre-canned" database schema suitable for storing data before inclusion into assessment models. It is not intended to be a primary source of data, rather a "staging area" for data to be gathered before any transformations are done to make the data suitable for e.g. a Gadget model.

A short tutorial on it's usage can be found as part of the online gadget course (https://gadget-framework.github.io/gadget-course/processing-input-data-with-mfdb.html).

- Supports both PostgreSQL or DuckDB databases
- No SQL knowledge required; the schema is created and managed automatically for you
- Supports storage of:
    - Area size / surface temperature, areas aggregated into divisions
    - Biological samples from survey / commercial fisheries
    - Stomach content surveys
    - Metadata on associated tows / vessels / bait type
- Easier model sharing:

  - Condense the source data required for a model run down to a single .duckdb file, or into a shared Postgres database, either of which are easy to share with other people
  - Consistent naming means less to figure out when receiving a model from someone else
  - Encourages automated scripting of model setup, which others can then modify and experiment with, e.g. changing length aggregation for a species
- Uses a "star schema" internally, for both efficiency and to check the correctness of input data

One major difference to generic database-querying tools is MFDB will preserve the aggregation levels used in formulating the results, which can then be used to automatically configure the model automatically using tools such as rgadget (https://gadget-framework.github.io/rgadget/) or read directly by gadget3 (https://gadget-framework.github.io/gadget3/).

MFDB can also be used to perform bootstrap sampling of data, as used in the Ling model described in (Elvarsson-2018)

MFDB is also used as a staging area for model output from the Atlantis ecosystem model, as part of the Fishing into the Future project (https://github.com/Fishing-into-the-future/atlantis_to_gadget). The process of using Atlantis as an operating model, ingesting it's output and using it to build an equivalent Gadget model, is scripted utilising MFDB. Thanks to this output from new Atlantis runs can be incorporated into Gadget with very little effort. The atlantis output files are read using an extension package mfdbatlantis (https://github.com/mareframe/mfdbatlantis).

Annex 4: Summary of presentations for ToR C

ToR c Simulated dataset and example model fit

Sarah Gaichas gave an overview of the simulated dataset for use in skill assessment in 2022, and expanded this dataset in 2023. Discussions on the use of the dataset continued using the mskeyrun GitHub discussion board in 2024. The work will be extended into WGSAM's next 3-year term.

Introduction

Ecosystem models can be complex, which can lead to high uncertainty in predictions (Hill *et al.* 2007). Skill assessment compares model predictions of interest with the truth from a system (Stow *et al.* 2009; Olsen *et al.* 2016). Understanding model skill can help us develop better models, as well as understand which models are most effective in which management situations. We wish to conduct skill assessment using simulated data because observations in fisheries are noisy, incomplete, and sometimes incorrect.

Simulating input data from an ecosystem model

We use existing Atlantis ecosystem model output to generate input datasets for a variety of multispecies models, so that the performance of these models can be evaluated against known (simulated) ecosystem dynamics. Atlantis models simulate a wide range of physical and ecological processes, include a full food web, and can be run using different climate forcing, fishing, and other scenarios.

We extract simulated data using the R package atlantisom. The purpose of atlantisom is to use existing Atlantis model output to generate input datasets for a variety of models, so that the performance of these models can be evaluated against known (simulated) ecosystem dynamics. Atlantis models can be run using different climate forcing, fishing, and other scenarios. Users of atlantisom will be able to specify fishery independent and fishery dependent sampling in space and time, as well as species-specific catchability, selectivty, and other observation processes for any Atlantis scenario. Internally consistent multispecies and ecosystem datasets with known observation error characteristics will be the atlantisom outputs, for use in individual model performance testing, comparing performance of alternative models, and performance testing of model ensembles against "true" Atlantis outputs.

The ms-keyrun simulated dataset

Our initial species selection includes 11 single species groups from the Norwegian Barents Sea (NOBA) Atlantis model (Hansen *et al.* 2016, 2019). These groups are fully age structured. All but two of them are fished.

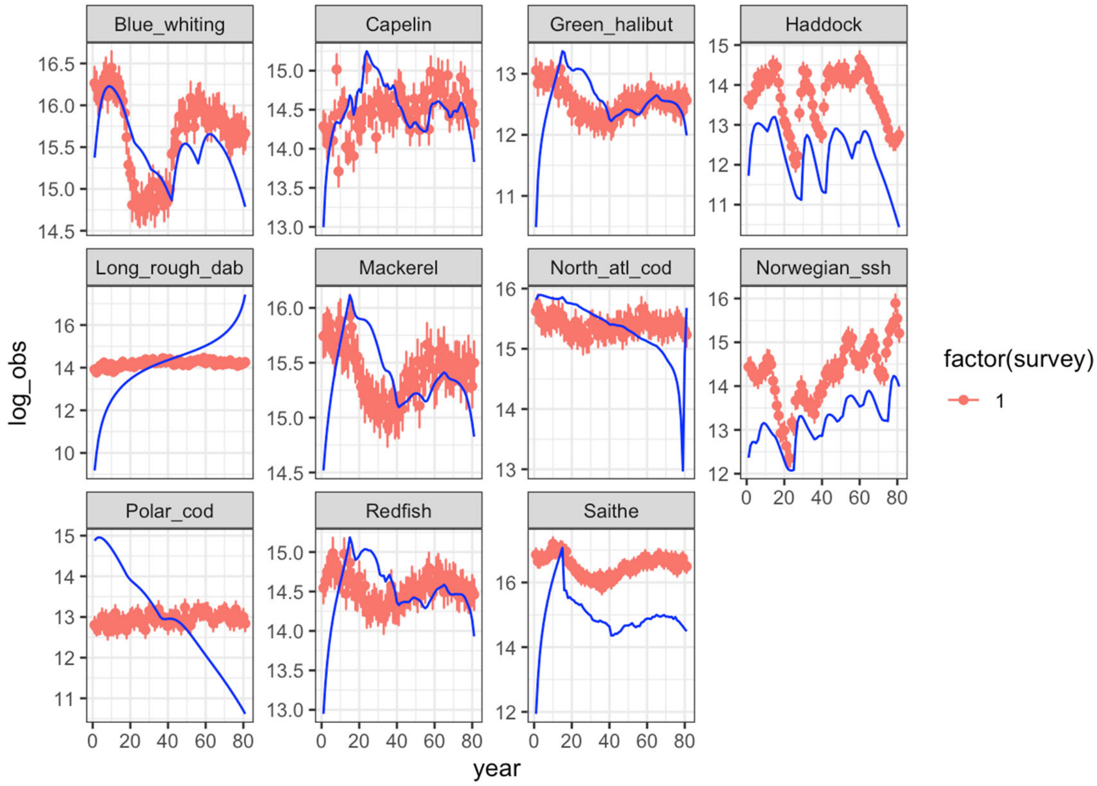
Table: Simulated species from NOBA Atlantis used for multispecies model testing

Model name	Full name	Latin name
Long_rough_dab	Long rough dab	*Hippoglossoides platessoides*
Green_halibut	Greenland halibut	*Reinhardtius hippoglossoides*
Mackerel	Mackerel	*Scomber scombrus*
Haddock	Haddock	*Melongrammus aeglefinus*
Saithe	Saithe	*Pollachius virens*
Redfish	Redfish	*Sebastes mentella*
Blue_whiting	Blue whiting	*Micromesistius poutassou*
Norwegian_ssh	Norwegian spring spawning herring	*Clupea harengus*
North_atl_cod	Northeast Atlantic cod	*Gadus morhua*
Polar_cod	Polar cod	*Boreogadus saida*
Capelin	Capelin	*Mallotus villosus*

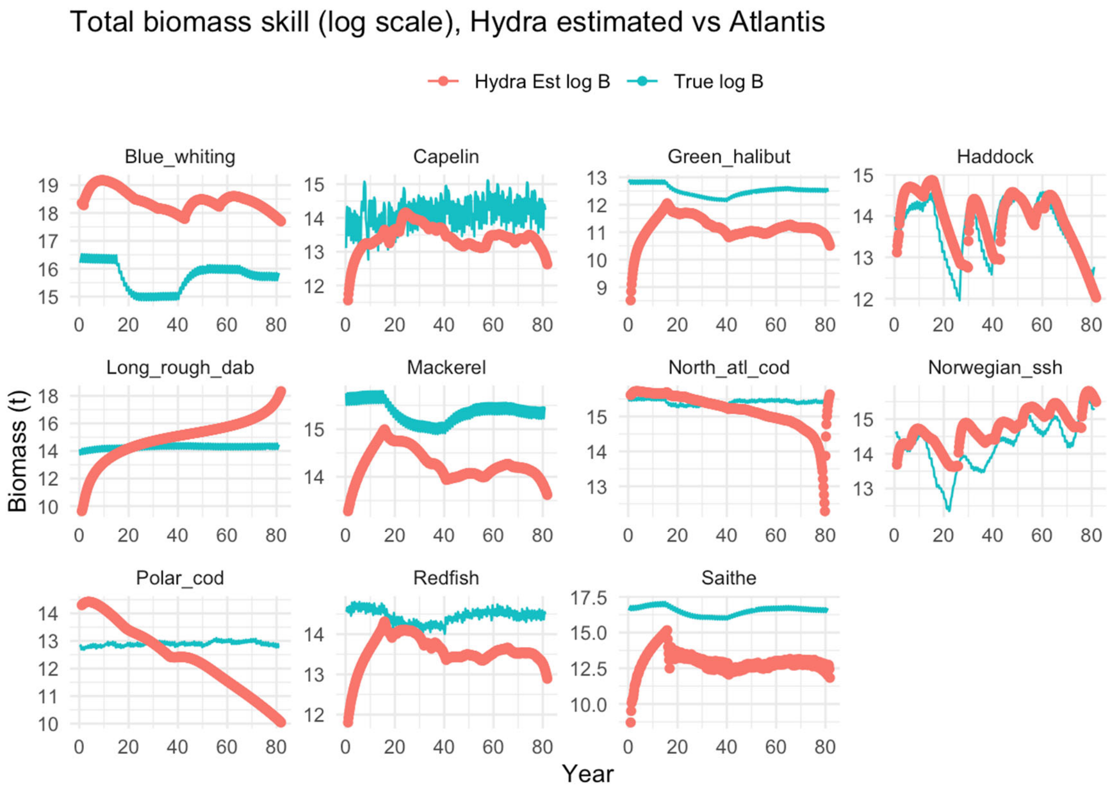
The full process for generating the simulated dataset is described on the mskeyrun R package documentation at this link.

Demonstration: Fitting a length-structured multispecies model to the data

The model is still in development, and *these are not final fits,* but we can compare not just fits to data but also skill against true Atlantis biomass, and calculate skill metrics. This can serve as an example for future collaborative work within WGSAM.



A Hydra fit



Hydra skill

References

Hansen, Cecilie, Kenneth F. Drinkwater, Anne Jähkel, Elizabeth A. Fulton, Rebecca Gorton, and Mette Skern-Mauritzen. 2019. "Sensitivity of the Norwegian and Barents Sea Atlantis End-to-End Ecosystem Model to Parameter Perturbations of Key Species." Edited by Athanassios C. Tsikliras. *PLOS ONE* 14 (2): e0210419. https://doi.org/10.1371/journal.pone.0210419.

Hansen, Cecilie, Mette Skern-Mauritzen, Gro van der Meeren, Anne Jähkel, and Kenneth F. Drinkwater. 2016. *Set-up of the Nordic and Barents Seas (NoBa) Atlantis Model*. Fisken Og Havet;2-2016. Havforskningsinstituttet. https://imr.brage.unit.no/imr-xmlui/bitstream/handle/11250/2408609/FoH_2-2016.pdf?sequence=1&isAllowed=y.

Hill, Simeon L., George M. Watters, André E. Punt, Murdoch K. McAllister, Corinne Le Quéré, and John Turner. 2007. "Model Uncertainty in the Ecosystem Approach to Fisheries." *Fish and Fisheries* 8 (4): 315–36. https://doi.org/10.1111/j.1467-2979.2007.00257.x.

Olsen, Erik, Gavin Fay, Sarah Gaichas, Robert Gamble, Sean Lucey, and Jason S. Link. 2016. "Ecosystem Model Skill Assessment. Yes We Can!" Edited by Carlo Nike Bianchi. *PLOS ONE* 11 (1): e0146467. https://doi.org/10.1371/journal.pone.0146467.

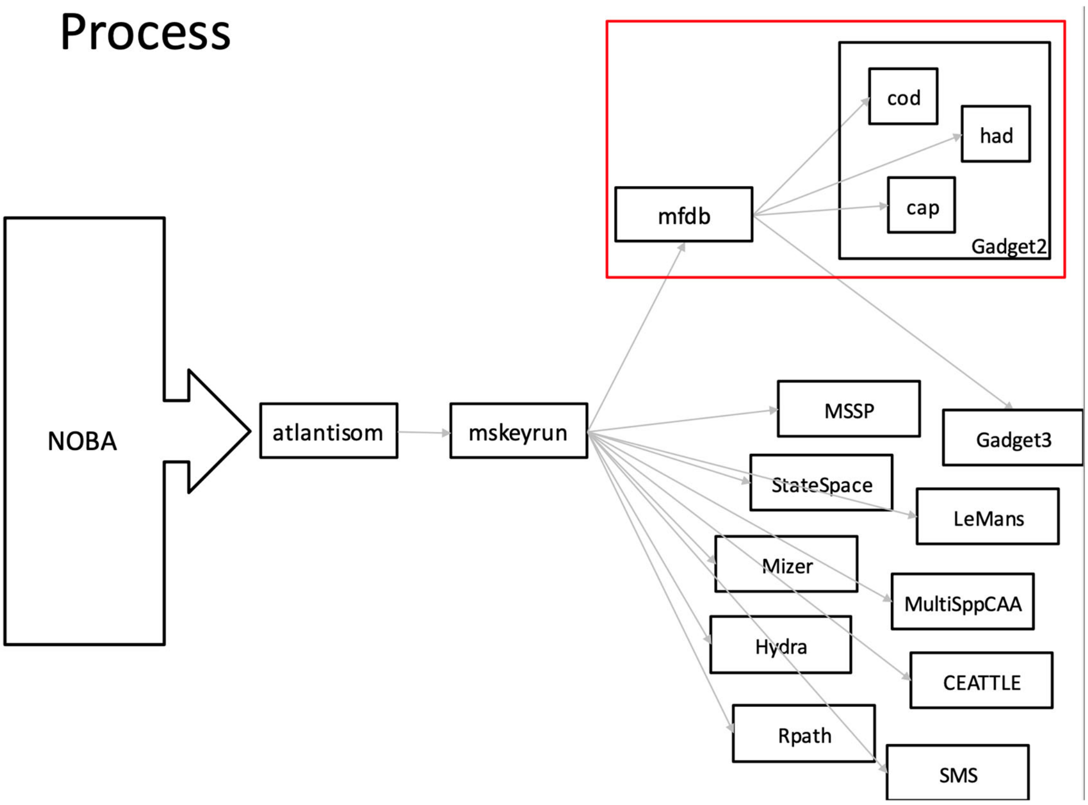
Stow, Craig A., Jason Jolliff, Dennis J. McGillicuddy, Scott C. Doney, J. Icarus Allen, Marjorie A. M. Friedrichs, Kenneth A. Rose, and Philip Wallhead. 2009. "Skill Assessment for Coupled Biological/Physical Models of Marine Systems." *Journal of Marine Systems*, Skill assessment for coupled biological/physical models of marine systems, 76 (1): 4–15. https://doi.org/10.1016/j.jmarsys.2008.03.011.

Summary of progress on fitting multi species models to simulated data

Gadget model

Valerio Bartolino, Alfonso Perez Bjarki Elvarsson Will Butler Daniel Howell

Overview of the process:



J dgjhw#vnlo#dvvhvvp hqw#z runiorz #

https://github.com/gadget-framework/wgsam-skill-assessment

mfdb

https://github.com/gadget-framework/mfdb

- organize OM simulated data to serve g2 and g3
- data already aggregated, so little use of aggregation procedures in this study - convenient way to build the model structure, model files and input data
- transparent and efficient, ie more simulations/scenarios will be available

Progress:

Three single species models (Northeast Atlantic cod (COD), Haddock (HAD), and Capelin (CAP)) have been implemented in Gadget2 using the simulated data. The pending issues for each stock are as follows:

HAD
- bound selectivity surveys
- estimate rec.len
- change rec to step 2 to capture rec peaks in the autumn survey - add both ALK

COD
- poor fit alk, strong residual pattern
- problem with weighting of the survey

CAP
- poor fit alk
- understocking (cap03)
- bound selectivity surveys
- marginally bound recruitment
- problem with weighting of the survey

The team plans to continue working towards a multispecies run once the individual single species performance is satisfactory. This is the typical Gadget workflow.

CEATTLE model

Grant Adams

Similar to Gadget, fitting of the CEATTLE age structured multi species model initially focused on single species fits to three species: Capelin, Haddock, Cod.

The models include both Spring and fall surveys, with 1 fishing fleet each, non-parametric time-invariant selectivity (fleet/season specific), analytical catchability (fleet/season specific)

Model code is at https://github.com/grantdadams/Rceattle and initial fit results are on the WGSAM GitHub Repo.

The model is fit with the following (data), likelihood components, and assumptions:

- Survey biomass (simSurveyIndex) : lognormal
- Catch (simCatchIndex): lognormal
- Fishery and survey age-comp (simXXXAgecomp) : multinomial
- Sample size? Used 200
- No length comp because avoiding "agecl" to "age" conversion
- Input weight-at-age (simXXXWtatAge)
- Maturity knife-edge at 0.5 max age
- Sex ratio = 0.5 for all ages
- M = 0.2 (estimating M hits bounds for capelin)

Overall, CEATTLE seems to be fitting capelin and haddock quite well. Cod is not fitting well (comp data are wonky and lack of trends in catch and index make it hard). Biomass and recruitment estimates seem reasonable?

Scripts developed to easily add in more species.

Next steps and recommendations include:

- Add in diet and bioenergetics data

- Still need to work through data to understand how to approach it
- Add in additional species
- Update cod data or remove from model
- Think through outline of intro/methods:
- I think we need a written framework to ensure simulation testing is well controlled
- Cant see true biomass estimates (otherwise well try and mimic them)
- What's our performance criteria (dictates what outputs I produce)

State-space model

Vanessa Trijoulet

The multi species state-space model converged, but didn't estimate uncertainties. This model produces similar outputs to CEATTLE, but with slightly less cod. Random walks were applied on recruitment. The model is able to follow the catch trends in the simulated data. The model was able to estimate F at age, but there is a need to adjust age as done for Gadget; Atlantis age 1 fish may be considered age 0 in an assessment context. The model was fit to total abundance from surveys. Age comp fits for cod were not great. Haddock fits were better, but fits were not great for capelin. Age comp in survey was terrible for cod, haddock not so bad, capelin bad in survey one but better in survey 2. Not age relabeling, may be an issue. Survey selectivity and fishing both showed a linear increase. Not tuned, just a test. Next steps are to analyze the issues raised above; it may help to relax selectivity.

Hydra model self tests

Cristina Perez

Hydra self testing: part of the review from WGSAM 2023.

Simulation testing for a subset of stocks for the Georges Bank ecosystem to understand the strength of trophic interactions, model performance and the ability to estimate key management quantities.

The model includes predation mortality on modeled prey and fits to: Stomach contents data, Survey and catch size compositions, Catch and survey abundance indices

Operating model: conditioned on the available data for two predator (Spiny dogfish, Atlantic cod) and two prey stocks (Atlantic mackerel, and Atlantic herring) and performed self-tests by fitting the model to data simulated from the operating model given no model misspecification.

Two different structural assumptions varying the parameter governing the amount of other food (OT) available

High values reducing the strength of the modeled species interactions

Smaller values increasing magnitude of predation mortality.

ToR c Mizer model skill assessment

Robert Thorpe gave a presentation on work carried out with Gustav Delius, Mike Spence, and Georg Engelhard using a mizer r-package for the North Sea under Tor C. The work was motivated by Mike Spence's and Michael Thompson's work on multi-model-multispecies MSY (MMMSY), where the constituent simulators (mechanistic models) project different yield curves and robustness of yield-mortality responses (Figure 1).

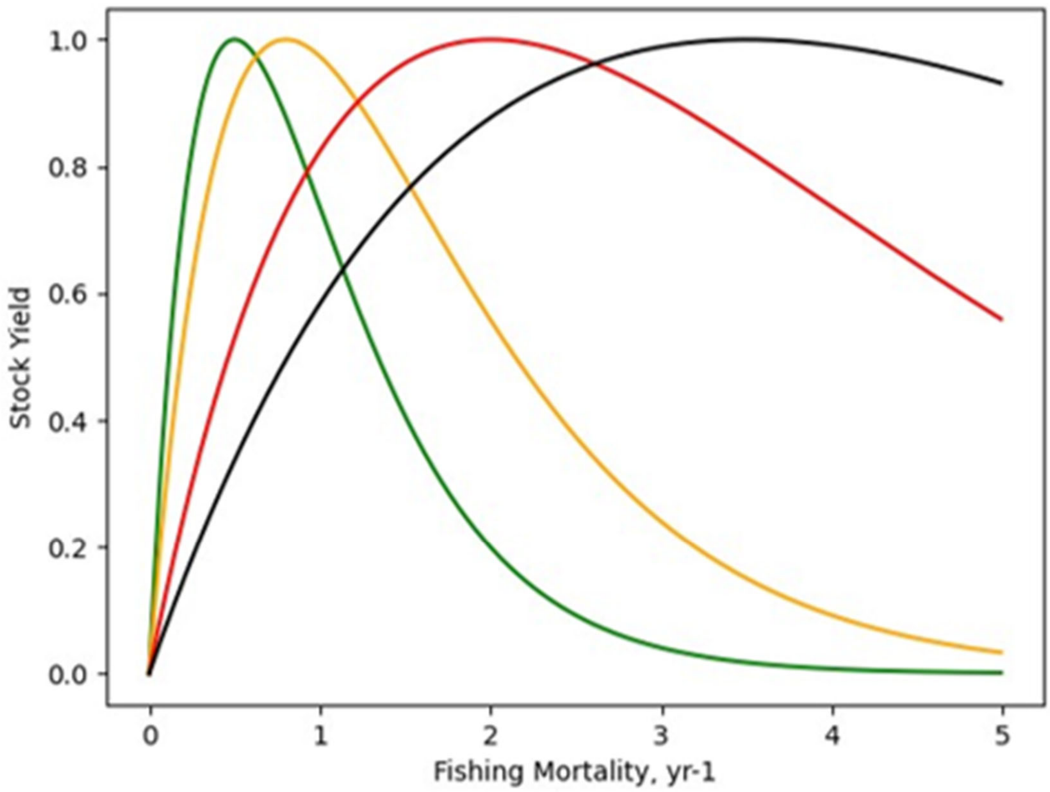


Figure 1: Schematic illustrating the different yield-mortality curve responses of the simulators (mechanistic models) in the Spence *et al.* (2018) ensemble MMMSY study. Green = EwE (Mackinson et el., 2018), orange = FishSUMS (Speirs *et al.*, 2014), red = Le Mans (Spence *et al.*, 2020; Thorpe *et al.*, 2015), black = mizer (Scott *et al.*, 2014).

The aim of this study was to see why mizer is so much more robust to high levels of fishing than EwE, and take steps towards understanding whether that feature is realistic. We did this by taking a "default" mizer from the r-package (Scott *et al.*, 2014), adding a number of stocks and setting their life-history parameters to be consistent with published LeMans models (Thorpe *et al.*, 2015: Thorpe and De Oliveira, 2018), and then simulating fishing mortality from 1880 until 2020 using estimates of fishing mortality reconstructed from fleet fishing segments. Foodweb interactions were then adjusted until projected biomasses adequately matched the 2010-2015 period. We subjected the model stocks to fishing mortalities ranging from F=0 to F=10 yr-1, using a single universal fleet with species-specific selection curves as per Thorpe *et al.,* 2015.

We found that the default mizer was very robust to fishing by the universal fleet, to an unrealistic extent. Reducing e-repro or the "evenness" of the foodweb (closeness to uniform 1s in the interaction matrix) led to less robust yield curves, but those approximating the shape of EwE only resulted when e-repro and foodweb evenness were both reduced, whilst the minimum individual size was increased, short circuiting some density dependence – Figure 2.

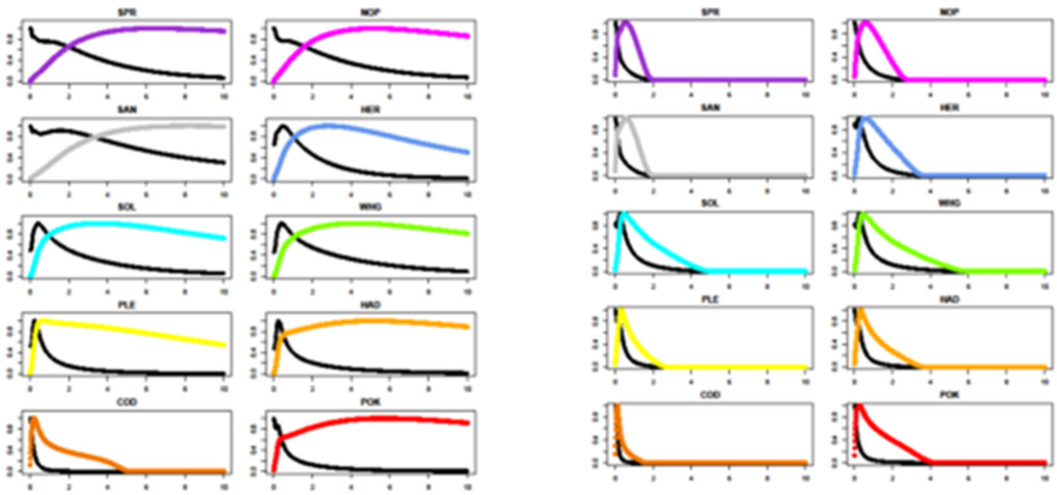


Figure 2: Fisheries yield curves for the default mizer (left) and following reductions to e-repro, foodweb evenness, and increases to the minimum individual size.

Both of these models can produce a skilful hindcast, despite having very different robustness to high levels of fishing, so we need access to addition sources of information to discriminate between them. We are currently working on using a hindcast of 20th century fishing to provide a skill assessment for the models. The rationale is that the longer time period can provide information on two regimes that have not existed in the recent past, firstly low levels of fishing, and secondly, a sudden resumption of fishing following a period of low fishing (immediately after WW1 and WW2). This may enable the feasible parameter space to be more constrained than hitherto.

Skill metrics fit into 3 main types – one set are general measures of generic outputs, (such as RMS error against 500mb geopotential height in meteorology, or stock biomasses in ecology), a second category are decision-related (e.g. keeping above limit biomasses), whilst a third relate to mechanistic responses (e.g not having a yield of >10% MSY for F>10). We are using all 3 types of skill metrics here for a selection of 100 mizer variants sampling possible parameter space. The assessment is ongoing, but so far we have found that 16 do not have significant decision-value, whilst all but 11 of the rest have fisheries yield responses that are too robust. These 11 remaining cover the full span of fidelity to biomass fits for 2010-15 (best and worst performing) suggesting that there is more work to do to establish the best-performing parameter space.

Other aspects of the 20th Century response were similarly problematic. The model fit to LPUE for plaice is good except before 1900, but for other demersal stocks it does not capture the recover between 1920 and 1960, and the modelled short-term response in the aftermath of the wars is poor, with the model wanting to recover stock sizes very quickly. The reasons for this remain to be discovered, but may include assumptions of scale-invariance, continuing density-dependence at small sizes, and assumptions about movement and prey selection, as well as treatment of the non-fish universe. Subsequent to the talk it was suggested that the assumption of continuous rather than pulsed recruitment for stocks in mizer made the modelled density dependence and stock resilience too high (Soudijn, 2016, pg 145).

References:

Mackinson *et al.*, 2018. Evaluating the fishery and ecological consequences of the proposed North Sea annual management plan. PLoS One, 13(1), e0190015. http://doi.org/10.1371/journal.pone.0190015

Scott *et al.*, 2014. Mizer: an R package for multispecies, trait-based and community size spectrum ecological modelling, Methods in Ecology and Evolution, 5(10), 1121-1125. https://doi.org/10.1111.2041-210X.12256

Soudijn F.H. 2016. Populations exposed to seasonal variability: from individual-level energetics to community dynamics, University of Amsterdam

Speirs *et al.* 2010. A length structured partial ecosystem model for cod in the North Sea. Fisheries Research, 106(3), 474-494: http:/doi.org/10.1016/j.fishres.2010.09.023

Spence *et al.*, 2018. A general framework for combining ecosystem models. Fish and Fisheries, 19(3), DOI 10.1111/faf.12310

Spence *et al.*, 2020. LeMaRns: A length-based multispecies analysis by numerical simulation in R, PLoS One, e0227767, DOI: 10.1371/journal.pone.00227767.

Thorpe *et al.*, 2015. Evaluation and management implications of uncertainty in a muti-species size-structured model of population and community responses to fishing. Methods in Ecology and Evolution, 6(1), 49-58, https://doi.org/10.1111/2041-210X.12292

Thorpe and De Oliveira, 2019. Comparing conceptual frameworks for a fish community MSY (FCMSY) using management strategy evaluation-an example from the North Sea, ICES Journal of Marine Science, 76(4), DOI. 10.1093/icesjms/fsz015

Calculating residuals for compositional data with correct properties

Vanessa Trijoulet, Christoffer Moesgaard Albertsen, Kasper Kristensen, Christopher M. Legault, Timothy J. Miller, Anders Nielsen

a) National Institute of Aquatic Resources, Technical University of Denmark, Kemitorvet 201, DK-2800 Kgs. Lyngby, Denmark

b) Northeast Fisheries Science Center, National Marine Fisheries Service, NOAA, 166 Water Street, Woods Hole, MA 02543, USA

Residuals are commonly used as diagnostics to validate assessment models. We demonstrate how common standardized residuals cannot be used for compositional data model validation and can lead to wrong conclusions due to correlation in the observations that are propagated into the residuals. We developed one-step-ahead (OSA) residuals for most commonly used multivariate distributions, i.e., multinomial, Dirichlet, Dirichlet-multinomial and logistic-normal. These residuals have the correct properties when the model is correct, i.e. they are independent (i), normally distributed (ii), with mean zero (iii), and with variance one (iv). We have developed an R-package “compResidual” (https://github.com/fishfollower/compResidual) and a Template Model Builder (TMB) contribution (https://github.com/vtrijoulet/OSA_multivariate_dists) that allow estimation of the OSA residuals externally to the model (models without random effects, compResidual) and inside the model (models with random effects, TMB contribution), respectively. These developments are relevant to any type of compositional data for single and multispecies models, i.e., aged- and length-structured data, stomach content data.

The paper is available in open-access at: https://www.doi.org/10.1016/j.fishres.2022.106487

The reference is as follows: Trijoulet, V., Albertsen, C.M., Kristensen, K., Legault C.M., Miller T.J., Nielsen, A. (2023). “Model validation for compositional data in stock assessment models: Calculating residuals with correct properties” Fisheries Research, Volume 257: 106487 DOI: 10.1016/j.fishres.2022.106487.

Annex 5: Summary of presentations for ToR D

Multi-species Maximum Sustainable Yield

Michael Thomson

Centre for Environment, Fisheries and Aquaculture Science

The UK Fisheries act 2020 defines maximum sustainable yield as the "highest theoretical equilibrium yield that can be continuously taken on average from a marine stock under existing environmental conditions without significantly affecting the reproduction process". This and other international agreements/legislation make it a requirement to manage fish stocks to MSY. Need a multispecies MSY (MMSY) as the yield of a species depends on fishing levels on other species. Considering species individually will give misleading results.

We define an MMSY using the concept of Nash equilibrium taken from the field of game theory. Defined as a strategy such that no player can improve their expected payoff by changing their strategy while every other player's strategy is unchanged. Applied to MMSY this means that changing the fishing mortality on any one species while leaving the fishing strategy on all other species unchanged does not improve yield.

We also need our fishing to be precautionary. This requires the long-term spawning stock biomass of each species to be above a certain limit such that the reproductive capability of the stock is not impaired with a set probability.

We need to predict the long-term yield and SSB of the system. Multiple mechanical ecosystem models exist for doing so. No one model is uniformly better than the others. Predictions and MMSY calculations are sensitive to the choice of model.

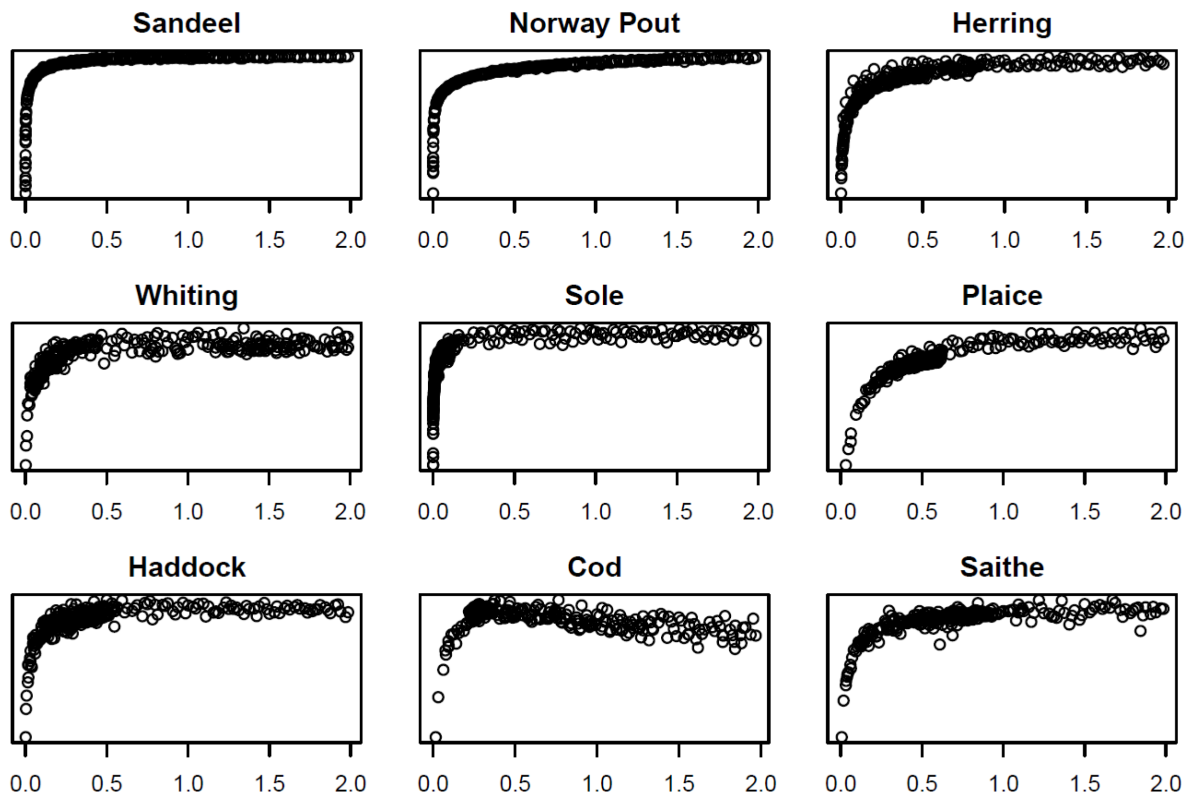
We combine predictions from multiple ecosystem models using an ensemble modelling approach. A statistical model defines relationships between the different ecosystem models the truth and observations from ICES stock assessments to produce a single prediction that combines data from all ecosystem models under consideration (see James Martindale's talk on EcoEnsemble for more details). Separate and independent ensemble models are fitting to yield and SSB.

To find the Nash equilibrium we need to be able to evaluate the long-term yield and SSB at all fishing mortalities. It is not computationally feasible to run the ecosystem models at all possible values. We fit separate Gaussian process emulators to each species under consideration to interpolate yield and SSB at fishing mortalities for which the ecosystem models have not been run.

We need an algorithm to find the Nash equilibrium. The currently proposed method iterates between each species under consideration searching for fishing mortalities that improve yield while satisfying MMSY and precautionary conditions.

We are studying the North Sea. Nine species are under consideration and four ecosystem models are being combine using the ensemble model for predictions.

Have obtained model runs for 334 fishing mortalities for each of the ecosystem models and presented long-term yield curves for each ecosystem model. As an example, these are the long-term yield curves for the ecosystem model Mizer.



How to use multiple marine ecosystem models (MEMs) to most appropriately and efficiently answer a complex management question?

Lehuta, S. Bourdaud, P., Girardin, R., Savina-Rolland, M, Vermard, Y., Marchal, P., Travers-Trolet, M.

Multi-model approaches are recognised as useful ways to deal with structural uncertainty in complex models, improve confidence in model results, at least when they converge, and provide better understanding of system functioning. These approaches require that the question at play is translated into a scenario that is implemented in several models ("one question, one scenario, several models"). Model results are then jointly analysed using model averaging or envelope technics. This approach is appropriate to evaluate stylised scenarios such as climate change projections with big drivers (IPCC, FishMIP). However, in the context of multi-criteria regional management evaluation, this approach is often unpractical.

These difficulties were experienced while trying to evaluate the impacts of the implementation of the landing obligation (LO) on the socio-ecosystem of the Eastern English Channel using three spatial ecosystem models: Atlantis (Girardin *et al.* 2018), Osmose (Bourdaud, 2018) and ISIS-Fish (Lehuta and Vermard, 2022). Although covering the same spatial footprint and fisheries, the three models made available to tackle this question within the European project DiscardLess, were fundamentally different in their structure and assumptions. They were originally built independently driven by different motivations. Their joint use was motivated by the availability of all modellers within the same institute and the need to present and justify this multiplicity of models to fisher's representatives, as partners of the European project.

The first attempt to use two of the models jointly, Atlantis and ISIS-Fish, to answer a question regarding selectivity improvements, rapidly evidenced the hurdles of the approach. Despite the simulation of the "same" output variables by the two models, plotting them on the same graph revealed inappropriate: On the one end because the way processes were modelled prevented scenarios to be exactly implemented the same way to the point that they needed to be considered two different scenarios; and on the other end, because of the way they were calibrated and run, namely at equilibrium after 100 years and dynamically over 10 years.

This failure pushed to elaborate and adopt a new approach of the use of multiple models that was developed and experimented within a EuroMarine project, M3E (Multiple Marine Ecosystem Models). The "multi-facet approach" was elaborated based on the case study of the impact of the LO in the Eastern English Channel by the Ifremer team and further completed and formalised during a workshop held online 11-15th of October 2021 with modellers from various European institutes (Travers-Trolet *et al.* in prep.). The basic paradigm could be summarised by "one question, several scenarios, several models" to point out the fact that the scenarios cannot be exactly similar across models but still can inform the same question. It advocated for a wise and restricted use of each model for its strength and purposes and for their combination through an interpretation of what they inform on, rather than a blind comparison of numbers. It attempted to remedy to common issues of multi-model exercises such as narrow scope, little operational use of scenarios, knocked-up scenarios, underutilisation/discard or inappropriate use of models.

It started with a conceptual cartography of the models' elements (species, fleets, management measures…), processes and simulations (time horizon, calibration…) with emphasis on the ecological meaning of the chosen implementation and elicitation of similarities and differences. Available scenarios were then dissected to identify the common question all models could inform, in this case the marginal effect on implementing the LO compared to a reference situation. The next phase was to list the different facets of the question possibly evaluated, and which model(s) could contribute to each facet.

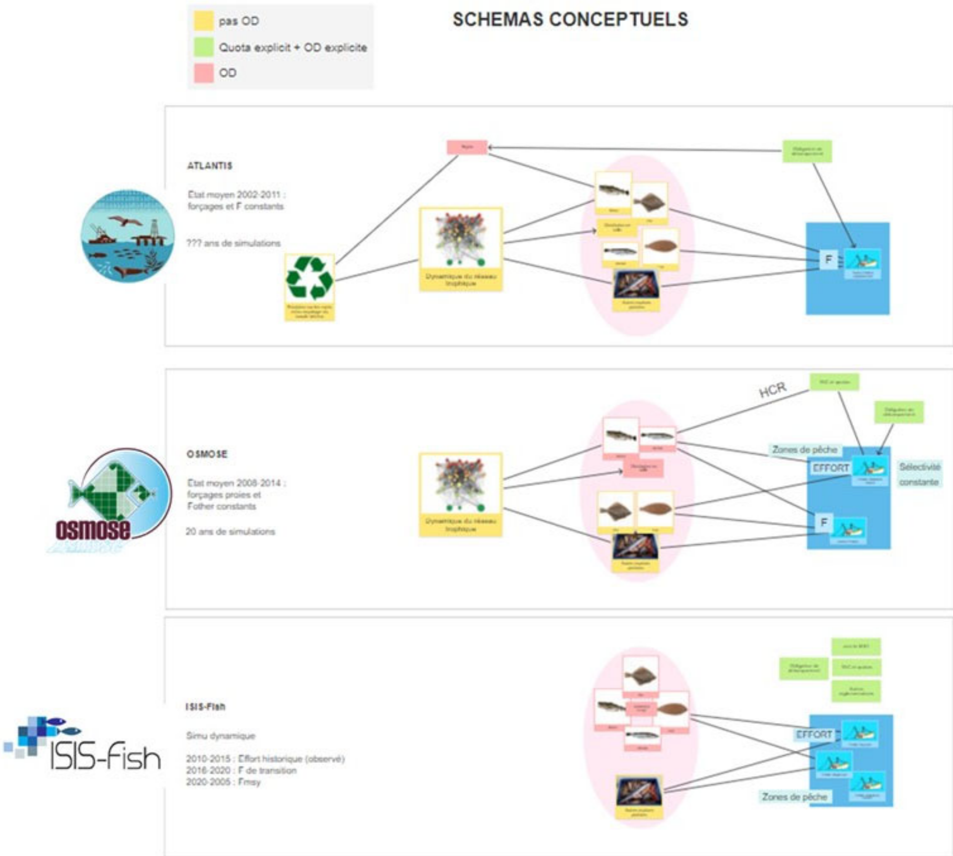


Figure 1: Conceptual mapping of models' elements.

The combination of Atlantis, Osmose and ISIS-Fish in the Eastern English Channel enabled informing the effects of the implementation of the LO on six facets of the question:

Facet	Models
Recycling and the quantity of food available to scavengers	Atlantis
Changes in fish predation	Atlantis and Osmose
Fisheries state variables (catch and biomass)	Atlantis, Osmose and ISIS-Fish
Mixed fisheries risks at the fleet level	ISIS-Fish
Fishers response	ISIS-Fish and Osmose

They fed a semi-quantitative story of these effects supported by either one or more of the models. When several models provided elements of response, differences were acknowledged and convergences were looked for, to evaluate which aspects were robust to scenarios and models. On a higher level, they evidenced the sensitivity of the answer to the processes modelled. For instance, the exercise established that effects on recycling were negligible and therefore it was acceptable to neglect them in Osmose and ISIS-Fish. On the contrary, changes in predation occurred and TACs were not always caught, pushing for an explicit modelling of predation and fishing mortalities. Jointly the three models covered the socio-ecological complexity of the management question and informed the multiple dimensions and criteria required by current management strategy evaluation demands.

References

Bourdaud, Pierre. « Impact of a landing obligation on coupled dynamics ecosystem-fishers: individual-based modelling approach applied to Eastern English Channel. PhD Thesis, Université du Littoral Côte d'Opale . https://archimer.ifremer.fr/doc/00440/55135/ ». PhD Thesis. Université du Littoral Côte d'Opale, 2018. https://archimer.ifremer.fr/doc/00440/55135/.

Girardin, Raphael, Elizabeth A. Fulton, Sigrid Lehuta, Marie Rolland, Olivier Thebaud, Morgane Travers-Trolet, Youen Vermard, et Paul Marchal. « Identification of the Main Processes Underlying Ecosystem Functioning in the Eastern English Channel, with a Focus on Flatfish Species, as Revealed through the Application of the Atlantis End-to-End Model ». Estuarine Coastal and Shelf Science 201 (5 février 2018): 208-22. https://doi.org/10.1016/j.ecss.2016.10.016.

Lehuta, S, et Youen Vermard. « Contrasting impacts of the landing obligation at fleet scale: impact assessment of mitigation scenarios in the Eastern English Channel ». ICES Journal of Marine Science, 27 septembre 2022, fsac148. https://doi.org/10.1093/icesjms/fsac148.

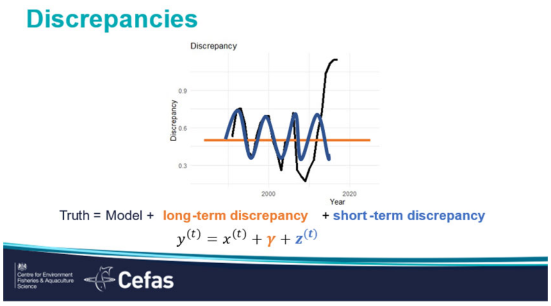
Articles in preparation about the multi-facet approach and its application to the Eastern English Chanel case study:

Travers-Trolet M., Piroddi, C., Romagnoni, G., Marchal, P., Shannon, L.J., Ortega, K., Bourdaud, P., Moullec, F., Spence, M.A., Poos, J.J., Thébaut, O., Martindale, J.A., Halouani, G., Savina-Rolland, M., Lehuta, S. Orchestrating scenarios for management of marine socio-ecological system: a new framework for multi-model approach. In. Prep.

Lehuta, S. Bourdaud, P., Girardin, R., Savina-Rolland, M, Vermard, Y., Marchal, P., Travers-Trolet, M. Multi-model insights on the impact of the LO in the Eastern English Channel: combine model responses to inform multiple facets of a complex management question. In prep.

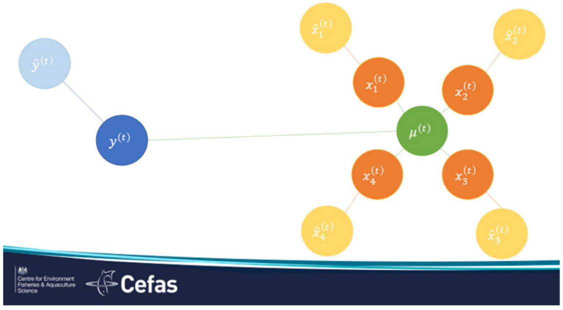
EcoEnsemble: James Martindale WGSAM Presentation
Ensemble modelling

- A 2018 paper by Spence et. al provides a general framework for combining multiple mechanistic models (simulators) into an ensemble while robustly quantifying the uncertainty.
- This approach involves statistically studying the discrepancies between the truth and the outputs of different simulators.
- Discrepancies are categorised into shared discrepancies, which are common to all the simulators, and individual discrepancies, which are specific to a given simulator. These discrepancies are further split into static long-term discrepancies and dynamic short-term discrepancies.



- We expect discrepancies to be similar for different simulators because these simulators are developed using the same scientific literature by people who often share ideas with one another. As such, we expect mechanistic simulators to share biases and have correlated discrepancy terms.

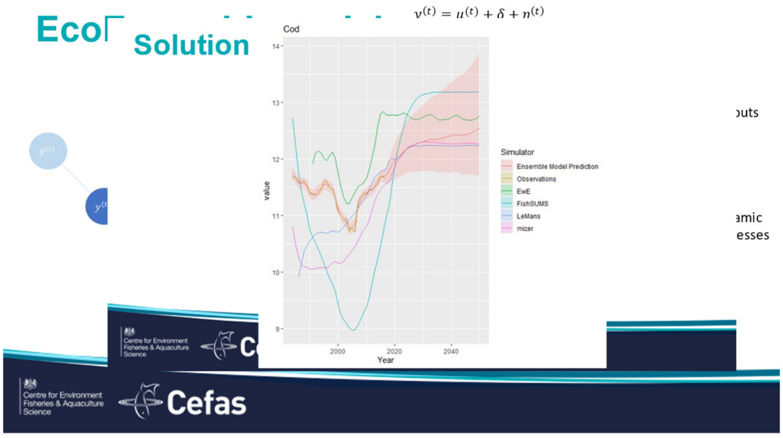
- The ensemble models the truth as a latent variable to be learnt, with additional latent variables for the "best guesses" of each of the simulators – the value of the simulators if there was zero parameter uncertainty, and all simulators produced all outputs for all



times.
- As we add simulators to the ensemble, we learn more about the distribution of discrepancies and our uncertainty of our overall prediction is reduced.

The EcoEnsemble package

- EcoEnsemble is an R package available on CRAN and GitHub to implement the ensemble modelling framework for a particular version of the ensemble model.
- In the EcoEnsemble model, the truth is modelled as a random walk and variables may either be available or missing for a particular simulator in the ensemble (i.e. the outputs may be missing some variables that are included elsewhere in the ensemble but all variables must be directly comparable – no aggregated sums of variables are sup-

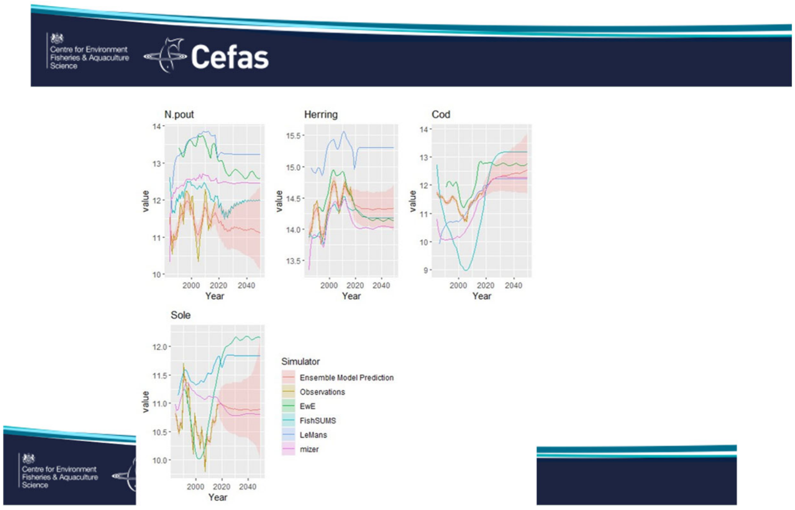


ported by the package).
- To use EcoEnsemble, priors for discrepancy terms may be manually configured, or default, uninformative priors are available.

- Functions exist to fit ensemble parameters to data, either by creating a point estimate that maximises the likelihood, or running an MCMC sampling using Stan. There also functions to generate samples of the latent variables (the truth and the simulator best guesses) and to plot the outputs of the ensemble.

Fitting the model

```
priors <- EnsemblePrior(4)
fit <- fit_ensemble_model(observations = list(SSB_obs, Sigma_obs),
                          simulators = list(list(SSB_ewe, Sigma_ewe, "EwE"),
                                            list(SSB_fs,  Sigma_fs, "FishSUMS"),
                                            list(SSB_lm,  Sigma_lm, "LeMans"),
                                            list(SSB_miz, Sigma_miz, "Mizer")),
                          priors = priors)
samples <- generate_sample(fit)
plot(samples)
```



Combining single-species models with simulation models

Michael A. Spence

Simulation models are potentially good at making predictions into the future, however, due to their flexibility they are unable to accurately predict the present or the short-term. On the other hand single-species models are have the flexibility to describe the present state as well as the short-term dynamics. In this presentation, I introduced an extension of the ensemble model of Spence *et al.* (2018), that uses a single-species model to determine the dynamics of the truth, along with simulation models. The ensemble model was fitted to catch at age and survey data for herring using four multispecies simulation models, describing their discrepancy. The aim is to be able to combine assessment and simulation models to make a formal framework across all time points.

Spence, MA, Blanchard, JL, Rossberg, AG, *et al.* A general framework for combining ecosystem models. Fish Fish. 2018; 19: 1031– 1042. https://doi.org/10.1111/faf.12310

Can we simultaneously adopt the precautionary approach for multiple species? A case study in the North Sea

Michael A. Spence

The precautionary approach to fisheries management requires taking account of uncertainty in the advisory process to ensure stock sustainability. Most fisheries management is based on a single-species approach, despite knowing that stocks interact through predation and competition for resources. The scale of these interactions is a response to the size of other stocks but are often treated as fixed relationships. So, a key question is `does the assumption of stock independence matter when setting reference points?' Here we examine the impact of species interactions on calculations of precautionary reference points using the North Sea as a case study. We combined four multispecies models using an ensemble model to robustly quantify the uncertainty and explore the fishing mortality that leads to nine species in the North Sea simultaneously following the precautionary approach. We found that the space of fishing mortality that leads to all species following the precautionary approach was much smaller than that of the individual species independently. Further, we found that zero fishing on all species does not mean that all the species follow the precautionary approach and some fishing on the larger species is required. We suggest that multispecies interactions should be considered when calculating precautionary reference points.

EcoEnsemble: A general framework for combining ecosystem models in R

Michael Spence

Often there are several complex ecosystem models available to address a specific question. However, structural differences, systematic discrepancies and uncertainties mean that they typically produce different outputs. Rather than selecting a single 'best' model, it is desirable to combine them to give a coherent answer to the question at hand. Many methods of combining ecosystem models assume that one of the models is exactly correct, which is unlikely to be the case. Furthermore, models may not be fitted to the same data, have the same outputs, nor be run for the same time period, making many common methods difficult to implement. In this paper, we use a statistical model to describe the relationship between the ecosystem models, prior beliefs and observations to make coherent predictions of the true state of the ecosystem with robust quantification of uncertainty. We introduce EcoEnsemble, an R package that takes advantage of the statistical model's structure to efficiently fit the ensemble model, either sampling from the posterior distribution or maximising the posterior density. We demonstrate EcoEnsemble by investigating what would happen to four fish species in the North Sea under future management scenarios. Although developed for applications in ecology, EcoEnsemble can be used to combine any group of mechanistic models, for example in climate modelling, epidemiology or biology.

There's more to ensemble models than model weights

Michael Spence

Often there are multiple models to describe a specific system and ultimately any decision can be sensitive to the model used. Choosing a single model from a suite of models is potentially throwing away information, meaning that the decision is not as well informed as it could be. Further, ignoring alternative models does not rigorously quantify the "true" uncertainty. An alternative to choosing a single model is to combine them using an ensemble model.

Many approaches to ensemble modelling involved weighting models and follow some strong assumptions, that a) are often not met, and/or b) are not utilising all the information. In this talk I will discuss these methods and compare them with alternative schemes. I will demonstrate that even in cases when the models are similar to the true data generating process alternative schemes do better than weighting the models.

In this talk I will discuss common ways of combining models, demonstrating them on a number examples and highlight that there is more to ensemble models than model weights.

TIME TRAVELLING IN THE NORTH SEA

Robert Thorpe gave a presentation on a 20th century reconstruction of North Sea fishing mortalities for 21 species carried out with Michael Spence, Michael Thomson, Georg Engelhard, and Gustav Delius. The reconstructions were generated from assessment data and unpublished estimates from John Pope, using mixed fisheries constraints and the EcoEnsemble package (Spence *et al.*, 2023) based on the Spence *et al.* (2018) ensemble method.

The reconstructions were then used as inputs to two fish community models (mizer – Scott *et al.*, 2014) and LeMaRns (Spence *et al.*, 2020) to make hindcasts of 20th Century landings which were compared with ICES (2023) landings data. The aim of the comparison was to untangle the effects of fishing and environmental change, given that the models assumed a constant environment, whilst the landings were influenced both by fishing and environmental change. Other potential benefits of the study were a) testing our understanding of short-term responses to rapid changes in fishing after the wars, demonstrating the utility of EcoEnsemble, and providing a long-term timeseries of fishing mortality that could be used as inputs to other models.

The experiment design is shown in Figure 1 (below):

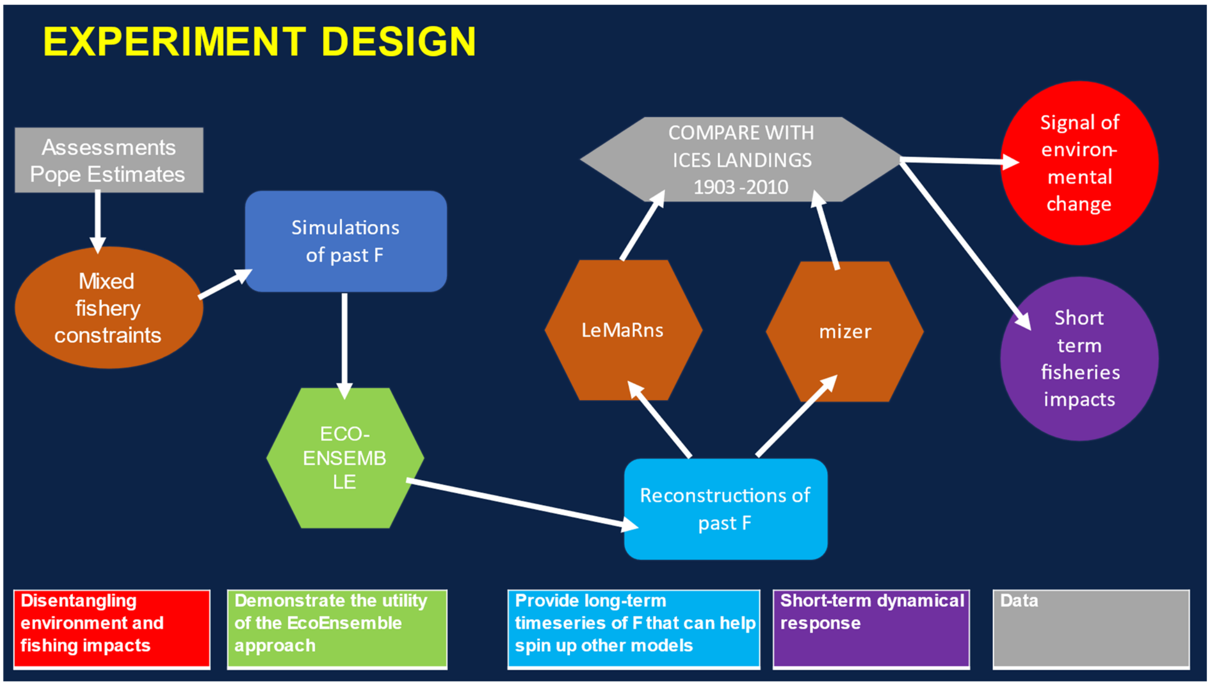


Figure 1: schematic of experiment design.

The reconstructed Fs from EcoEnsemble take account of biases relative to assessment data and are systematically different from the mixed fishery estimates as shown in Figure 2 (magenta line), with lower uncertainty, particularly early in the 20th century, when the mixed fishery estimates were relatively unconstrained. This reflects the ability of EcoEnsemble to maximise the available information. Other benefits include reduced bias and improved fidelity of model hindcasts driven by the EcoEnsemble estimates of F.

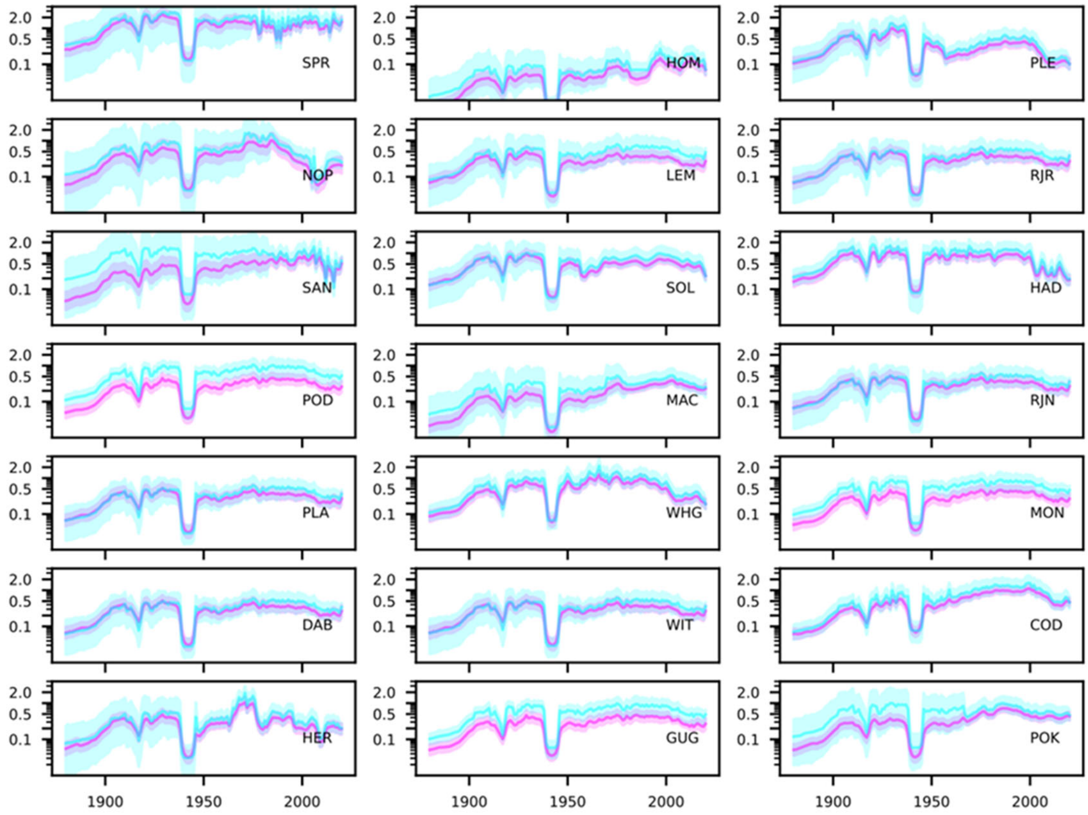


Figure 2: Mixed fishery (blue) and EcoEnsemble (magenta) estimates of fishing mortality (F) on a log scale.

Hindcasts made using the EcoEnsemble F timeseries and assuming a constant environment appear to suggest 3 environmental regimes, as shown in Figure 3.

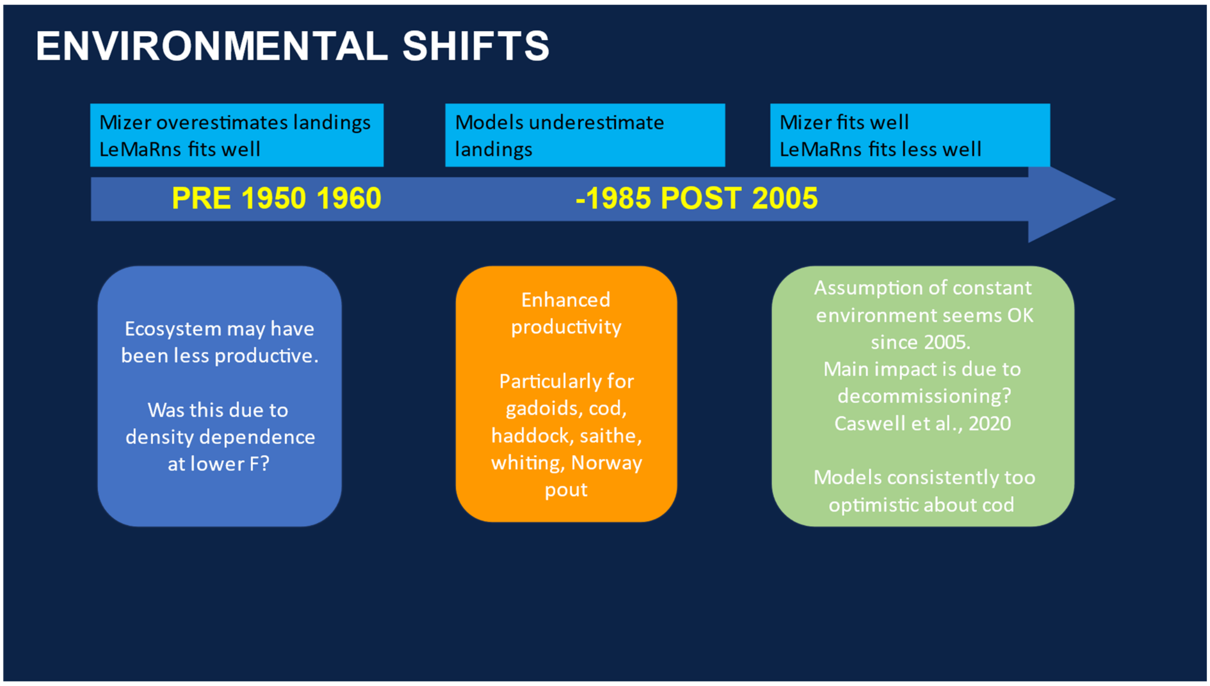


Figure 3: Suggested environmental regimes from comparison of the mizer and LeMaRns hindcasts of ICES landings.

We conclude that climate and fishing are both important drivers of stock abundance in the North Sea since 1900. The simulations also provide valuable insights into short-term dynamics in the periods after WW1 and WW2, and a timeseries of F that can be used to help spin up other models, as well as demonstrating the utility of the EcoEnsemble approach.

References:

ICES. 2023. Catch statistics. https://www.ices.dk/data/dataset-collections/Pages/Fish-catch-and-stock-assessment.aspx

Scott, F, Blanchard, JL, Andersen, KH. 2014. Mizer: an R package for multispecies, trait-based and community size spectrum ecological modelling, Methods in Ecology and Evolution, 5(10), 1121-1125

Spence, M. A., Bannister, H. J., Ball, J. E., Dolder, P. J., Griffiths, C. A., and Thorpe, R. B. 2020. LeMaRns: A Length-based Multi-species analysis by numerical simulation in R. PLOS ONE, 15(2), 1–12

Spence, M. A., Blanchard, J. A., Rossberg, A. G., Heath, M. R., Heymans, J. J., Mackinson, S., *et al.* (2018). A general framework for combining ecosystem models. Fish Fish. 19, 1031–1042. doi: 10.1111/faf.12310

Spence, MA, Martindale, JA, Thomson, MJ. 2023. EcoEnsemble: A general framework for combining ecosystem models in R. DOI 10.1111/2041-210X.14148.

FishMIP update – Tyler Eddy

FishMIP (Fisheries and Marine Ecosystem Model Intercomparison Project) aims to understand the impacts of climate change on marine ecosystems. FishMIP is a network of over 100 global and regional marine ecosystem modellers developing standardized protocols for projecting climate change impacts. The models used include a variety of approaches, such as size or age-based, food-web, and species distribution models. FishMIP's objectives include improving model reliability, promoting climate-resilient food security, protecting biodiversity, and testing climate solutions. Ongoing research within FishMIP focuses on refining regional modelling and aligning global models with regional scales. Tools like the FishMIP Input Explorer app and new data resources (e.g., World Ocean Atlas) support this work. Future scenarios, based on socio-economic pathways (SSPs) and oceanic system pathways (OSPs), are being developed to project the impacts of fishing and climate change under various conditions, guiding adaptation strategies and conservation efforts. More information available at fishmip.org.

Newfoundland & Labrador ecosystem modelling update Species Distribution Modelling (SDM): Research by Ruiz-Díaz *et al.* focuses on modelling how climate change will impact the distribution of snow crab and Atlantic cod. These models provide projections on how the species' habitats will shift under future climate conditions.

Ecosystem Roles of Marine Species: We are exploring the roles of different species in Newfoundland and Labrador ecosystems. For example, the role of harp seals and northern shrimp in the food web is being examined, as well as how phytoplankton productivity and capelin abundance affect copepod populations. We are also exploring the role of zooplankton in two-way coupled ecosystem models.

Developing EcoEnsemble: A Simulation Study

Michael Thomson

Centre for Environment, Fisheries and Aquaculture Science

EcoEnsemble (Spence, Martindale, & Thomson, 2023) is an R package for implementation of a Bayesian statistical ensemble model. The ensemble model combines predictions from multiple "simulators" (eg. Multispecies fisheries models) into a single predictor of some variables of interest (VoI). The relationship between the simulators and the VoI is described using some discrepancy terms. These discrepancies are decomposed into short and long-term and shared and individual parts. So there are 4 types of discrepancy, individual short-term, individual long-term, shared short-term, shared long-term. There is one of each type of shared discrepancy which is shared between all simulators while each simulator has its own individual discrepancies.

To fit the model EcoEnsemble uses the no-U-turn sampler as implemented in Stan (Stan Development Team, 2023). To use Stan we need to define priors for the discrepancy terms. This requires placing priors on several covariance matrices. EcoEnsemble decomposes these covariance matrices into correlation and variance matrices. EcoEnsemble provides several choices of prior distribution, LKJ, inverse-Wishart, the method of concordance and hierarchical.

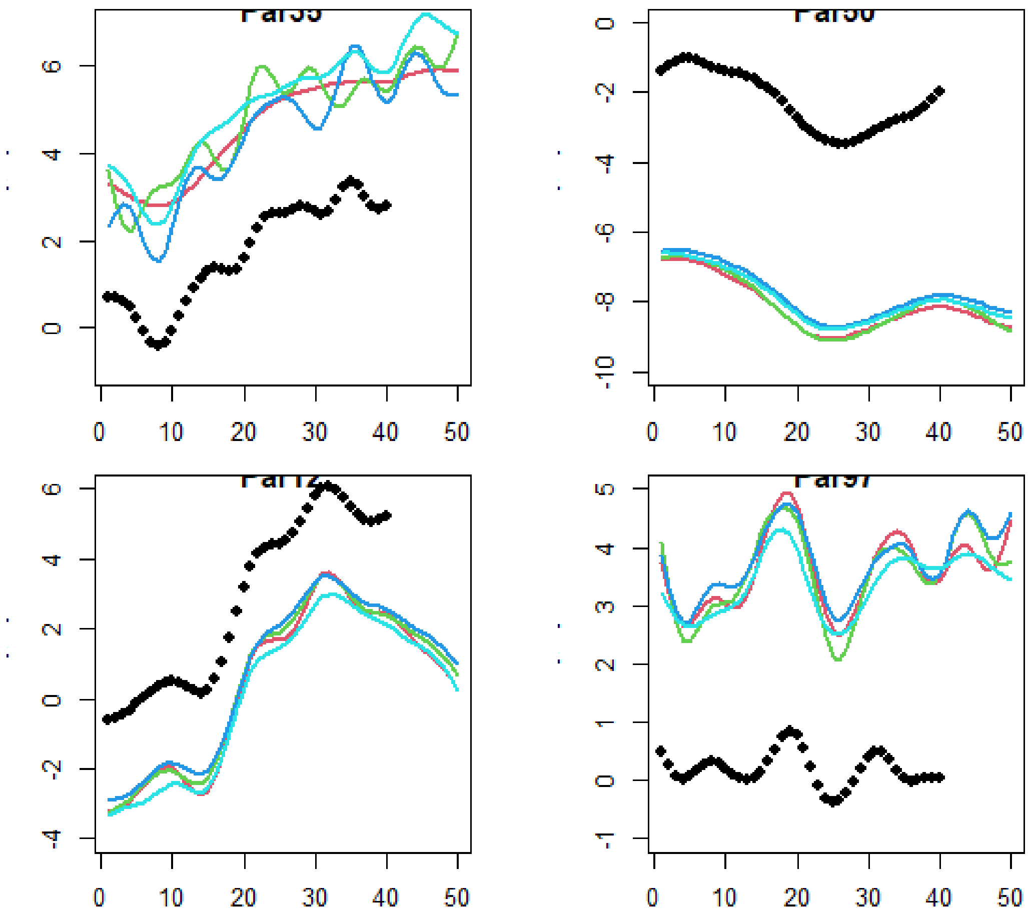
EcoEnsemble includes an example dataset with four species from the North Sea. We used this to test the effect of prior choice. As a default prior on the correlation matrices we used an LKJ(1) prior as it is uniform on the space of correlation matrices. However, this required very tight priors on the variance matrices (except for the shared long-term).

The ensemble model requires high-dimensional priors including objects with complex structure (correlation matrices) which are hard to visualise. Often high-dimensional, complex priors can lead to unintended structure being introduced into a prior. This has motivated a simulation study where we can test priors applied to synthetic data where we know the truth to allow us to provide advice to users of the model and construct sensible default priors.

Correlation matrices have complex structures and it's difficult to visualise a distribution over the space of correlation matrices. We use an archived R package (Tokuda, Goodrich, Van Mechelen, Gelman, & Tuerlinckx, 2009) to produce a number of plots that allow us to examine properties of distributions of covariance and correlation matrices.

EcoEnsemble's inverse-Wishart prior takes the unusual step of restricting the inverse-Wishart distribution, usually distributed over covariance matrices, and restricting it to covariance matrices. Using the visualisation package we are able to identify that whereas with the standard inverse-Wishart distribution increasing the degrees of freedom parameter leads to correlations concentrating around 0, the correlation matrix restricted inverse-Wishart has the opposite behaviour. Correlations become more strongly concentrated at the boundary values, -1 and 1.

Given a prior the ensemble model is fully defined to allow us to simulate from that prior to produce a synthetic data set consisting of some set of synthetic simulators and synthetic observed data. Applying that to the prior with tight variance priors and LKJ(1) priors on the correlation matrices we see that this prior is making a strong assumption. Plots (see below) show that for almost every draw from this prior the simulators follow the short-term behaviour of the truth very closely while being displaced from the truth by some amount. This amount is very similar for each simulator. This is most likely because the prior for the shared long-term variance is much greater than those of the other discrepancy terms and so it dominates. Prior simulation has thus allowed us to identify that this prior is in fact making a strong assumption about the relation of the simulators to each other and the truth.



References:

Spence, M. A., Martindale, J. A., & Thomson, M. J. (2023). EcoEnsemble: A general framework for combining ecosystem models in R. *Methods in ecology and evolution*, *14*(8), 2011-2018.

Stan Development Team. (2023). Stan Modelling Language Users Guide and Reference Manual 2.33. Retrieved from https://mc-stan.org

Tokuda, T., Goodrich, B., Van Mechelen, I., Gelman, A., & Tuerlinckx, F. (2009). *Visualising distributions of covariance matrices*. Retrieved from http://www.stat.columbia.edu/~gelman/research/unpublished/Visualization.pdf

Annex 6: Summary of presentations for ToR E

Assessing resource resilience in relation to fisheries objectives

Valerio Bartolino, Andrea Belgrano

Department of Aquatic Resources, Swedish University of Agricultural Sciences (SLU)

Understanding the level of uncertainty and buffers required in fishery advice and management is central aspect of a precautionary approach to fisheries management. Increasing changes in climate call urgently for a better understanding of the short-, medium and long-term impacts on fish stocks and require evaluation of the robustness of the current management targets in relation to raising uncertainties. In addition to a progressive warming detected for many sea basins, climate change is expected to increase the intensity and frequency of extreme events. Hence, the current management to be precautionary must guarantee stocks resistance to the occurrence of extreme events (hereafter referred as shocks), and the potential for stock recovery.

A simulation study was set up to evaluate the present F_{MSY} framework for the management of cod, herring and sprat in the Baltic Sea. A Multispecies age-length based model was used to reconstruct the population dynamics of the Eastern Baltic cod (cod.27.24-32), Baltic sprat (spr.27.22-32) and Central Baltic herring (her.27.25-2932) stocks (hereafter referred as cod, herring and sprat). The model is built using the statistical multispecies modelling framework Gadget https://github.com/gadget-framework/gadget2 and it has a particular focus to reconstruct the dynamics of these three stocks while explicitly accounting for predation mortality caused by cod on the clupeids. The model used in this simulation study is an extension of the model presented at WGSAM (ICES 2019) which was already extending the model initially presented by Kulatska *et al.* (2019).

Ten-years simulations were performed forced by environmental conditions expected for the period 2021-2030. Uncertainty on the projected stock trajectories under the different climate scenarios was represented in two main ways: (1) as structural uncertainty in the climate projections and (2) as stochastic deviations of recruitment from a segmented stock-recruitment relationship with an environmental driver.

Plausible, ecosystem coherent climate scenarios were derived by extraction of the different hydrographic drivers of fish dynamics across different warming (RCP4.5 and RCP8.5) and nutrient scenarios from RCO-SCBI provided by the Swedish Metereological Institute (Fig. 1; Saraiva *et al.* 2019). This biogeochemical model provided coherence in the projected physical environment used in this study. Uncertainty on future climate projections is incorporated by the use of three climate models. The use of coherent environmental scenarios via a coupled physical-biogeochemical model makes difficult to generate extreme events (i.e., heatwave) coherent with the rest of the bio-physical system in a predefined moment in time. For this reason, we used coherent environmental scenarios to drive long-term changes in the stocks (i.e., recruitment), while the effects of a short-term shock were simplified in this study by simulating recruitment events at the limits of the distribution of recruitments observed (i.e., recruitment failure corresponding to the lowest recruitment observed in the whole time series).

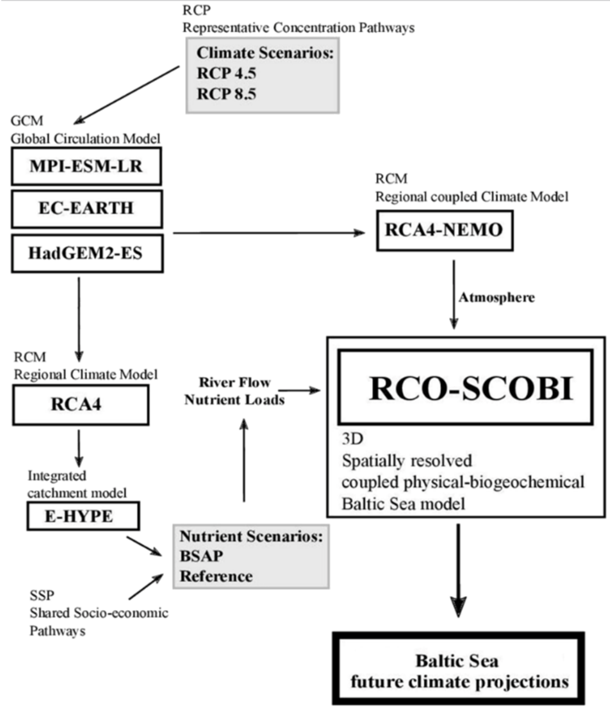


Figure 1. Conceptual diagram of the modeling framework providing hydrographic projections for the Baltic Sea (modified from Saraiva *et al.* 2019).

Future recruitment projections for the three species were based on environmental sensitive stock-recruitment relationships parametrised on the historical time period using Gadget SSB-Recruitment estimates and selected hydrographic variables. In the specific, the cod reproductive volume and summer temperatures were used as external drivers of the recruitment of cod and the two clupeids. A similar simple model formulation was adopted for the three stocks where deviations from a hockey-stick stock-recruitment model were explained by a linear relationship with stock specific hydrographic variables and a log-normal distributed error term. 1000 simulations of recruitment were generated for each scenario (fig. 2). Three shock scenarios were implemented: no shock, one shock in the first year of the simulation period, and two shocks assumed in the first year of the simulation and another randomly during the ten years of the simulation.

Cod - RV	Herring – SST78	Sprat – SST8

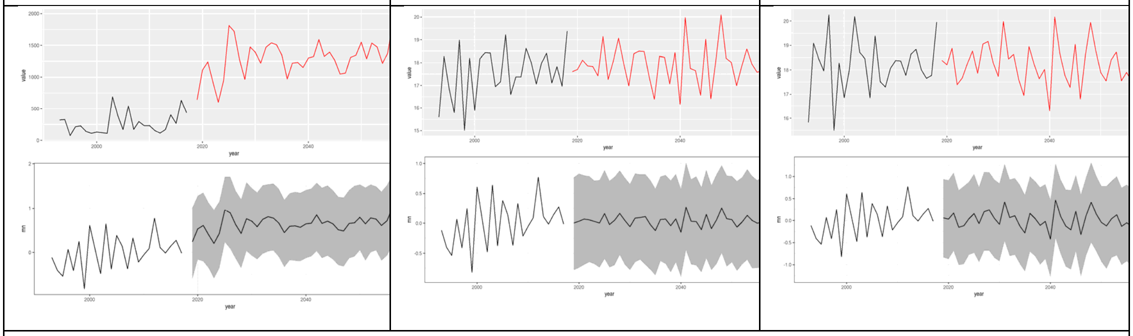


Figure 2. (top) time series of key environmental drivers from one of the GCM models (EC-EARTH) and scenarios (RCP4.5 – BSAP). (bottom) recruitment log-deviations from S-R segmented regressions for cod, herring and sprat illustrated for this specific scenario and climate model.

Two fishing level scenarios were evaluated for herring and sprat based on harvest rates which approximate the F_{MSY} and F_{MSY} lower range values. For cod, only one fishing scenario was implemented corresponding to a low recent exploitation level (average of 2016-2018).

Fish stocks response to the shock(s) and the performance of exploitation levels corresponding to F_{MSY} and F_{MSY} lower range were evaluated in terms of resistance (i.e., ability of the stocks to oppose to the perturbation) and resilience (i.e., ability of the stocks to recover from the perturbation). Five metrics were adopted:

- Amplitude - ratio between the minimum SSB reached in response of the 2020 shock and the SSB in 2020
- Responsiveness - number of years between the shock and the minimum observed stock level
- Biological risk - probability of SSB falling below $B_{trigger}$ after the shock in the period 2021-2030
- Recovery rate - probability that stock level is at $B_{trigger}$ or above in 2030
- Recovery speed - number of years to reach stock levels corresponding to MSY after the SSB drop from the 2020 shock

Within the conditions and limits of these simulations we found that:

- The positive effect on clupeids recruitment of increasing temperature within the range predicted for the coming decade was more visible on sprat stock dynamics while it was nearly negligible for herring
- Cod recovery was expected to be very slow even under a low F, and reduced eutrophication as in the Baltic Sea Action Plan was a prerequisite
- Accordingly, cod predation on clupeids increased in the ten years of the simulation but slowly
- Exploitation levels within the MSY framework [F_{MSY} – F_{MSY} low] were likely to keep both herring and sprat above $B_{trigger}$ also in the light of occasional poor recruitment so were found to be compatible with the objective of building resilience
- A more precautionary exploitation level set at the lower bound of the F_{MSY} range generated a visible buffer against unaccounted uncertainties. Simulations showed that while this might be unnecessary for sprat it could be more relevant for herring

References

ICES. 2019. Working Group on Multispecies Assessment Methods (WGSAM). ICES Scientific Reports. 1:91. 320 pp.

Kulatska N., Neuenfeldt S., Beier U., Elvarsson B. Þ., Wennhage H., Stefansson G., Bartolino V. 2019. Understanding ontogenetic and temporal variability of Eastern Baltic cod diet using a multispecies model and stomach data. Fisheries Research, 211: 338–349.

Saraiva S., Meier H.E.M., Andersson H., Höglund A., Dieterich C., Gröger M., Hordoir R., Eilola K. 2019. Uncertainties in Projections of the Baltic Sea Ecosystem Driven by an Ensemble of Global Climate Models. Front. Earth Sci. 6: 244.

New England EBFM prototype MSE

Sarah Gaichas

NMFS Northeast Fisheries Science Center, National Oceanic and Atmosperic Administration, Woods Hole, USA (NOAA)

1. Overview

Combining existing tools to evaluate EBFM procedures: Link multispecies model to MSE framework with built in single species assessment tools for New England prototype EBFM MSE (pMSE)

Fundamental objectives:

- Maintain or increase inflation adjusted total value for the fishery
- Preserve ecosystem function and structure
- Maintain stock complex biomass around levels that optimize fishing opportunities
- Prevent overfishing
- Response of regulations to stocks at low abundance, and recovery of depleted stocks
- Reduce regulatory complexity
- Example explanatory materials, including Rshiny for exploring results
- Food web model Rpath MSE capabilities
- Atlantis potential

2. EBFM pMSE

Hydra: multispecies operating model conditioned on Georges Bank data within MSE framework as prototype test of EBFM strategies.

Results: additional flexibility and increased yield possible with EBFM "ceilings and floors" without increased risk to single stocks.

3. Operating model building blocks
    a) Multispecies catch at length Hydra (Gaichas *et al.*, 2017)

Hydra-Associated GitHub repositories

- hydra-sim (Simulation Model Wiki): https://github.com/NOAA-EDAB/hydra_sim/wiki
- hydra-sim (estimation fork): https://github.com/thefaylab/hydra_sim
- hydradata (estimation fork): https://github.com/thefaylab/hydradata
- hydra-diag (diagnostics): https://github.com/thefaylab/hydra_diag

    b) Groundfish MSE framework (Mazur *et al.*, 2023)

Groundfish MSE repository, branches, and tools

- main branch (including wiki): https://github.com/lkerr/groundfish-MSE
- pMSE branch: https://github.com/lkerr/groundfish-MSE/tree/JJ-EBFM
- includes ASAP model and Plan B smooth
- uses SAMtool state space surplus production assessment

4. Setup
    a) Management Alternatives

- Single-species management, with stock-specific assessments and catch advice.

- Single-species management, with stock-specific assessments and catch advice and dynamic reference points.
- Stock complex management, with stock-complex level assessments and abundance index thresholds.
- Stock complex management, with stock-complex level assessments and abundance index thresholds, with gear-based stock complexes.
- Stock complex management, with stock-complex level assessments and dynamic reference points.

b) Operating Model Scenarios

- Base
- Fleet dynamics (Adjust q for economic value)
- Initial biomass (Below base)
- Predator pressure (Increase M1 <30cm)
- Prey change (Increase other food)

5. Outreach materials

a) Summary pages

- pMSE Overview
- Management Objectives and Performance Metrics
- Management Alternatives
- Results Engagement
- Operating Models
- Assessment Models

b) Rshiny https://gavinfay.shinyapps.io/results-viewer/

References

Gaichas, S. K. *et al.* (2017). "Combining stock, multispecies, and ecosystem level fishery objectives within an operational management procedure: simulations to start the conversation". In: *ICES Journal of Marine Science* 74.2, pp. 552-565. ISSN: 1054-3139. DOI: 10.1093/icesjms/fsw119. URL: https://academic.oup.com/icesjms/article/74/2/552/2669545/Combining-stock-multispecies-and-ecosystem-level (visited on Oct. 18, 2017).

Mazur, M. D. *et al.* (2023). "Consequences of ignoring climate impacts on New England groundfish stock assessment and management". In: *Fisheries Research* 262, p. 106652. ISSN: 0165-7836. DOI: https://doi.org/10.1016/j.fishres.2023.106652. URL: https://www.sciencedirect.com/science/article/pii/S0165783623000450.

ISIS-Fish as the operating model of a Management Strategy Evaluation framework applied to common sole (*Solea solea*) in the Bay of Biscay

HERNVANN Pierre-Yves[1], LEHUTA Sigrid[1], MAHEVAS Stéphanie[2], PELLET Ian[1], RICOUARD Antoine[2], LECOMTE Jean-Baptiste[1], LE PAPE Olivier[3]

[1] DECOD (Ecosystem Dynamics and Sustainability), Institut Agro—Agrocampus Ouest, IFREMER, INRAE, 44311 Nantes, France

[2] UMR 9190 MARBEC, University of Montpellier-IRD-IFREMER-CNRS, Av. Jean Monnet, CS 30171, Sète Codex 34203, France

[1] DECOD (Ecosystem Dynamics and Sustainability), Institut Agro—Agrocampus Ouest, IFREMER, INRAE, 35042 Rennes, France

Common sole (Solea solea) of the Bay of Biscay (BoB) is a commercially important and emblematic stock whose sustainable exploitation is challenged by both the highly mixed aspect of the demersal fisheries exploiting it, and changes in productivity observed over the recent decades. The presentation given by the authors at the WGSAM 2024 is relevant with the WG ToR e, dedicated to management strategy evaluation (MSE). It introduces a MSE framework using as an operating model a fisheries dynamics model, i.e. ISIS-Fish. Unlike most models discussed at the WGSAM meeting, ISIS-Fish does not represent species interactions (e.g. trophic). However, by accureately depicting mixed fisheries and their technic interactions, it can simulate the reciprocal effects of one specie's dynamics and its management on fisheries dynamics, and their indirect impacts on other species' status and harvest levels.

The present framework consists of a MSE loop that connects a sole stock assessment model (XSA model) used by the ICES WGBIE, to an operating model, i.e. the spatially-explicit fisheries dynamics model ISIS-Fish. ISIS-Fish represents fisheries dynamics through three modules interacting together: one biological module dynamically modelling fish multiple fish populations (age-, length- structured or biomass pool), one human module describing the structure of fleets demersal and pelagic fleets through their technical characteristics, gears and strategies, and one module mimicking management strategies. Thus, the MSE loop allows to represent (i) the impact of mixed fisheries on sole population while integrating (ii) the environmental drivers of the latter's productivity, (iii) the management procedure from data collection to the recommendation of a TAC based on perceived biomass and (iv) the consequences of new management target for fleet dynamics and population status …and so on.

In order to represent the effect of environment on sole productivity within the MSE framework, we used a nursery habitat suitability model for sole integrating river flow as a covariate. The MSE framework is then used to test the robustness of alternative management procedures regarding changes in river flow. The efficiency of these management procedures are compared to that of the one currently in place, i.e. the MSY approach. A first set of procedures consisted of two versions of harvest control rules, more or less cautious, that scale up or down the maximum fishing pressure allowed according to river flow anomalies recently observed. In a second type of management procedure, the MSY harvest control rule is kept but recruitment anomalies predicted by the nursery habitat suitability model are incorporated in short-term recruitment projections used by the stock assessment to generate the catch advice. The analysis of performance metrics among simulated management procedures suggests that all alternative management procedures integrating the environmental signal efficiently contribute to stock rebuilding. While the most efficient regarding this aspect are the environmentally-informed harvest control rules, they also bring more inter-annual variability in total catch (hence fishing revenue), while this variability is lower when the management procedure is based on an environmentally-informed stock assessment.

Preliminary though, the outcomes this novel MSE framework are promising for a transition to a climate-informed EBFM that does not forget human dimensions. While only basic stock-status related performance metrics have been explored in these preliminary results, other indicators will be assessing the indirect effects of single-species designed management procedures for sole on the status of other species. More importantly, it will also allow quantifying the consequences of these procedures for the Bay of Biscay fisheries while accounting for heterogeneity in their characteristics and technical interactions among them, in particular via the calculation of socio-economic indicators. In the next steps, his framework will integrate for through to 2100. The integration of environment-productivity relationships and the coupling to stock assessment models could also be extended to other species represented in the model.

MULTISPECIES MANAGEMENT AND THE COMMON FISHERIES POLICY (CFP)

Robert Thorpe gave a presentation on work carried out with Michael Spence and Matthew Kerr, which uses a multi-model ensemble study of around 20,000 simulated outcomes for fishing nine stocks to see how different interpretations of multispecies MSY would be expected to perform.

Article 2(2) of the CFP states that "The CFP shall apply the precautionary approach to fisheries management and shall aim to ensure that exploitation of living marine biological resources restore and maintains populations of harvested species above levels which can produce the maximum sustainable yield.". This gives rise to four key aspects of interpretation, the standard ICES approach to which is given in table 1.

Table 1: Key issues concerning the meaning of CFP Article 2(2) and the standard ICES interpretations of them.

Issue	ICES interpretation
What is the unit of management?	The single species (stock) in most cases.
What is the precautionary approach?	B is above Bpa such that the probability that B is in fact below Blim is less than 0.05.
What is the maximum sustainable yield?	Maximum stock yield such that B is greater than Bpa, assuming a constant environment.
What do we mean by keeping populations above levels that can produce MSY?	B is above Bpa. This can be justified by the use of a hockey stick stock recruit relation, such that if B> Blim the stock recruitment is at a maximum.

However, these interpretations become problematic when multispecies interactions are considered, because the precautionary space will shrink with increasing numbers of interacting stocks if defined in this way. In this study we use an app designed by Michael Spence and Matthew Kerr (Spence and Kerr, 2023) to look at the expected SSB, fisheries yields, and risks of stock depletion (below Blim or relative to unfished biomass) for nearly 20,000 fisheries simulations for 9 North Sea stocks for which estimate of Blim were available (sandeel, Norway pout, herring, whiting, sole, plaice, haddock, cod, and saithe) in order to see how some alternative definitions of these concepts would perform. A schematic of the study is shown in Figure 1.

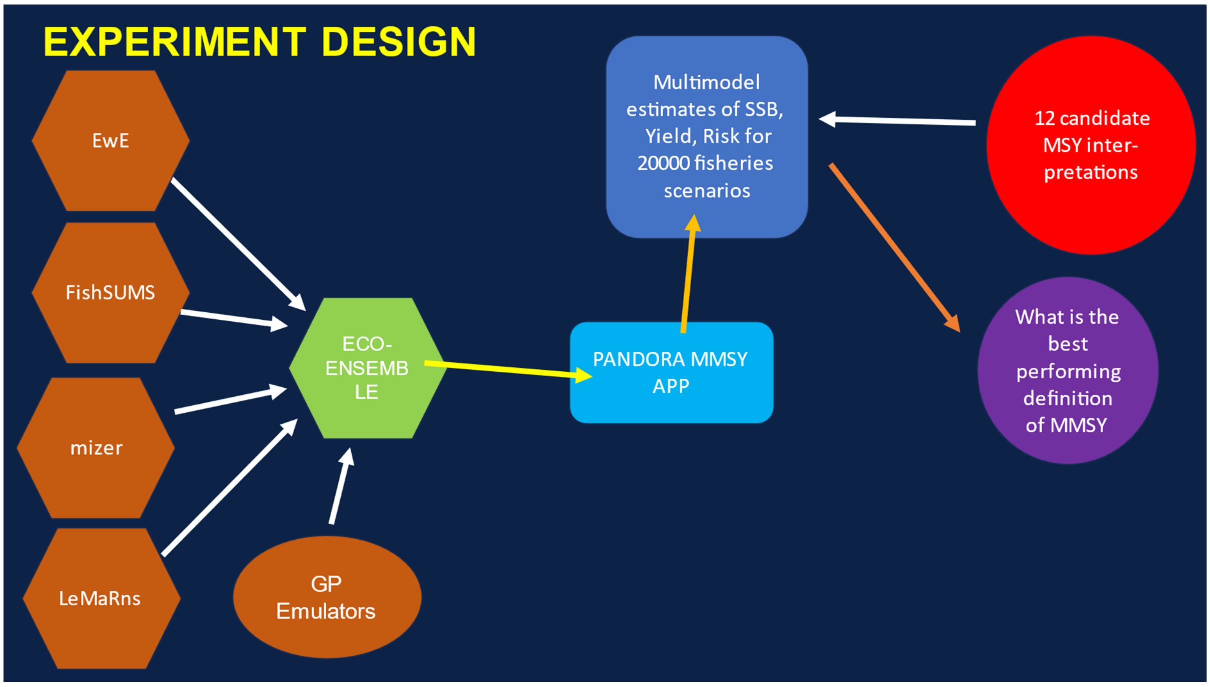


Figure 1: Schematic of the multispecies MSY in the context of the CFP study.

We considered 4 different ways of evaluating the risks of stock depletion across the 9 stocks and 3 different ways of determining the optimum combination of Fs, covering a range of 12 interpretations for more single species to more multispecies in nature. Outcomes expressed in expected levels of yield and risk are shown in Figure 2.

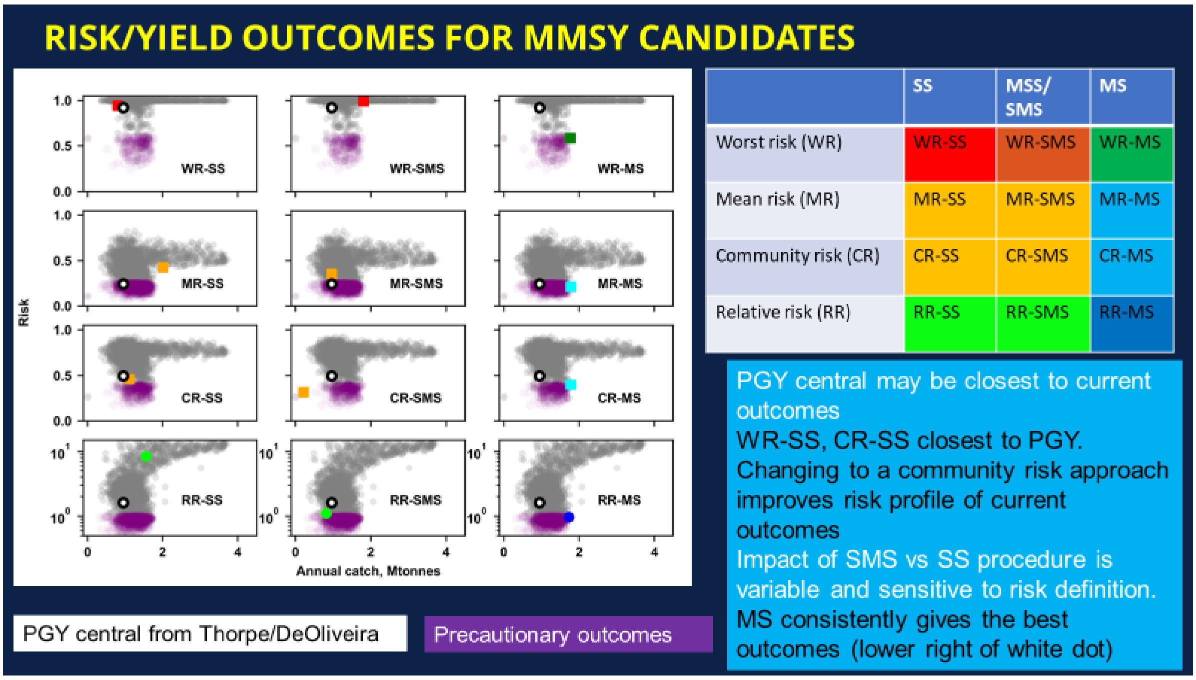


Figure 2: Expected long-term risk and reward outcomes for 12 possible interpretations of Article 2(2). Colours show the range of interpretations from more like single species (red) to community-wide (blue). The precautionary space as defined by the interpretation is purple.

We found that overall the multispecies interpretations performed better, provided that we could know enough about the management space to define them sufficiently. The study showed that the levels of risk generally were much higher than those explicitly acknowledged, as a result of a) structural uncertainty in modelling the multispecies community, and b) the incommensurate nature of the set of Blims in the context of the whole community, and we need to be honest with stakeholders about the higher risks that are involved in stock management.