ICES WGSAM REPORT 2018

HUMAN ACTIVITIES, PRESSURES AND IMPACTS STEERING GROUP

ICES CM 2018/HAPISG:20

REF. ACOM, SCICOM

Report of the Working Group on Multispecies Assessment Methods (WGSAM)

15–19 October 2018

Paris, France

International Council for the Exploration of the Sea

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

Recommended format for purposes of citation:

ICES. 2018. Report of the Working Group on Multispecies Assessment Methods (WGSAM), 15–19 October 2018, Paris, France. ICES CM 2018/HAPISG:20. 89 pp.

Contents

Executive summary **3**

1 Administrative details **4**

2 Terms of Reference **4**

3 Summary of Work plan **5**

4 Summary of Achievements of the WG during 3-year term **5**

5 Final report on ToRs, workplan and Science Implementation Plan **5**

5.1 ToR A. Report on further progress and key updates in multispecies and ecosystem modelling throughout the ICES region 5

5.2 ToR B. Update of key-runs (standardized model runs updated with recent data, producing agreed output and agreed upon by WGSAM participants) of multi-species and ecosystem models for different ICES regions (Baltic Sea EwE, LeMans Framework proposed for use in Irish Sea) 7

5.3 ToR C. Consider methods to assess the skill of multispecies models intended for operational advice 12

5.4 ToR D. Investigate the performance of multi-model ensemble in comparison to single model approach 13

5.5 ToR E. Test performance and sensitivity of ecosystem indicators 14

5.6 ToR F. Metanalysis of impact of top predators on fish stocks in ICES waters 16

5.7 ToR G. Explore the consequence of multispecies, mixed fisheries interactions and environmental factors in practical multispecies advice for fisheries management (MSY related and other biological reference points) 17

6 Cooperation **22**

7 Summary of Working Group self-evaluation and conclusions **22**

Annex 1: List of participants **23**

Annex 2: Recommendations **25**

Annex 3: WGSAM resolution 2019–2021 .... **26**

Annex 4: WGSAM self-evaluation 2016–2018 .... **28**

Annex 5: ToR A. Review further progress and deliver key updates in multispecies and ecosystem modelling throughout the ICES region **32**

Annex 6: ToR C. Consider methods to assess the skill of multispecies models intended for operational advice **58**

Annex 7: ToR D. Investigate the performance of multi-model ensemble in comparison to single model approach **63**

Annex 8: ToR E. Test performance and sensitivity of ecosystem indicators............70

Annex 9: ToR F. Metanalysis of impact of top predators on fish stocks in ICES waters...71

Annex 10: ToR G. Explore the consequence of multispecies, mixed fisheries interactions and environmental factors in practical multispecies advice for fisheries management...73

Executive summary

The Working Group on Multispecies Assessment Methods (WGSAM) has held three meetings between 2016 and 2018, chaired by Sarah Gaichas, USA, and Alexander Kempf, Germany.

WGSAM is continuously making significant contributions to enable ICES to develop its capability in giving model-based multispecies and ecosystem advice, in particular, on the ecosystem impacts of fishing and climate change. This is a priority area identified in the ICES strategic plan and is consistent with scientific needs to support implementation of the Common Fisheries Policy and Marine Strategy Framework Directive.

A particularly important contribution by WGSAM were the guidelines on quality assurance of multispecies and ecosystem models intended for advice. WGSAM members have further developed and demonstrated methods for sensitivity analysis and skill assessment of complex models that can be adopted by the wider community; and are collaborating with Working Group on Integrative, Physical-biological and Ecosystem Modelling (WGIPEM) on a best practices paper. WGSAM has also developed and presented different approaches for estimating multispecies and ecosystem based reference points and tested them within multiple simulation frameworks. Finally, major progress was made on operationalizing management strategy evaluation (or structured decision-making) integrating both quantitative modelling and stakeholder engagement processes to address key management questions.

WGSAM executed key runs for the Baltic Sea (Ecopath with Ecosim in 2016) and the North Sea (SMS in 2017). WGSAM has also worked with WGIPEM and WGMIXFISH to identify joint prioritises for further developments.

1 Administrative details

Working Group name
Working Group on Multispecies Assessment Methods (WGSAM)
Year of Appointment within current cycle
2016
Reporting year within current cycle (1, 2 or 3)
3
Chair(s)
Sarah Gaichas, USA
Alexander Kempf, Germany
Meeting venue(s) and dates
Reykjavik, Iceland, 10–14 October 2016 (17 participants)
San Sebastian, Spain, 16–20 October 2017 (21 participants)
Paris, France, 15–19 October 2018 (16 participants)

2 Terms of Reference

ToR A. Review further progress and deliver key updates in multispecies and eco-system modelling throughout the ICES region.

ToR B. Update of key-runs (standardized model runs updated with recent data, producing agreed output and agreed upon by WGSAM participants) of multi-species and ecosystem models for different ICES regions (Baltic Sea EwE, LeMans Framework proposed for use in Irish Sea).

ToR C. Consider methods to assess the skill of multispecies models intended for operational advice.

ToR D. Investigate the performance of multi-model ensemble in comparison to single model approach.

ToR E. Test performance and sensitivity of ecosystem indicators.

ToR F. Metanalysis of impact of top predators on fish stocks in ICES waters.

ToR G. Explore the consequence of multispecies, mixed fisheries interactions and environmental factors in practical multispecies advice for fisheries management (MSY related and other biological reference points).

3 Summary of Work plan

Year 1	Work on all ToRs. Focus on ToR e, f and g. ToR b: Keyruns (as required)
Year 2	Work on all tors. Focus on ToR c and d. ToR b: Keyruns (North Sea SMS, as required)
Year 3	Work on all tors. Focus on Synthesis ToR c-g. ToR b: Keyruns (as required)

4 Summary of Achievements of the WG during 3-year term

- Key runs: Modelling outpus and advisory products
  - Baltic EwE (see 2016 report), used in structured decision making (see ToR g, this report)
  - North Sea SMS (see 2017 report)
- M values from SMS are used in stock assessments in the North Sea area: Cod, haddock, whiting, herring, sprat and sandeel.
- Model framework reviews
  - LeMans Ensemble (see 2016 report)
    - Used in WKIRISH Advisory process
  - FLBEIA (see 2017 report)
  - Multispecies state-space model (see 2017 report)
- Methodological developments
  - Towards best practices for model skill assessment (ToR c)
    - Sensitivity analysis in complex models (see 2016–17 reports and this report, ToR c)
    - Effect of diet data quality/quantity on multispecies model performance (2018 ToR c)
  - Initial work on multi-model comparisons and ensemble methods for complex models (ToR d, all years)
  - Estimation of multispecies fishery reference points (ToR g, all years)
  - Evaluation of performance of different reference points (ToR g, all years)
- Papers reported in Section 5 below under Scientific contributions.

5 Final report on ToRs, workplan and Science Implementation Plan

Progress and fulfilment by ToR, with science highlights by ToR

5.1 ToR A. Report on further progress and key updates in multispecies and ecosystem modelling throughout the ICES region

WGSAM received updates over the course of 2016–2018 on the Greenland and Iceland Seas, Barents Sea, Norwegian Sea, Celtic Sea, North Sea, Baltic Sea, South European Atlantic Shelf, and US Northwest Atlantic.

Yearly reporting current progress of multispecies models and predator-prey research in ICES Ecoregions, noting in particular:

1) Continued development of existing established modelling approaches such as Stochastic Multispecies Model, Gadget, Ecopath with Ecosim.
2) Progress made with the development of ATLANTIS ecosystem models in Iceland, in major updates to the Northeast US Atlantis, and continued development in other regions.
3) The surge in development of multiple models capable of simulating multispecies and ecosystem processes across regions.
4) The use of multi-model approaches that include multispecies and ecosystem models to provide insight into management-relevant parameters and processes.

Relevant Papers

Sturludottir, E., Desjardins, C., Elvarsson, B., Fulton, E. A., Gorton, R., Logemann, K., and Stefansson, G. 2018. End-to-end model of Icelandic waters using the Atlantis framework: Exploring system dynamics and model reliability. Fisheries Research, 207: 9–24.

Buchheister, A., Miller, T. J., Houde, E. D. (2017a). Evaluating ecosystem-based reference points for Atlantic menhaden (Brevoortia tyrannus). Marine and Coastal Fisheries. doi: 10.1080/19425120.2017.1360420.

Buchheister, A., Miller, T. J., Houde, E. D., and Loewensteiner, D. A. (2017b). Technical Documentation of the Northwest Atlantic Continental Shelf (NWACS) Ecosystem Model. Report to the Lenfest Ocean Program, Washington, D.C. University of Maryland Center for Environmental Sciences Report TS-694–17. Available at:

http://hjort.cbl.umces.edu/NWACS/TS_694_17_NWACS_Model_Documentation.pdf .

Busch, D. Shallin, and Paul McElhany. 2016. “Estimates of the Direct Effect of Seawater pH on the Survival Rate of Species Groups in the California Current Ecosystem.” *PloS One* 11 (8): e0160669.

Hodgson, Emma E., Isaac C. Kaplan, Kristin N. Marshall, Jerry Leonard, Timothy E. Essington, Shallin D. Busch, Elizabeth A. Fulton, Chris J. Harvey, Albert Hermann, and Paul McElhany. 2018. “Consequences of Spatially Variable Ocean Acidification in the California Current: Lower pH Drives Strongest Declines in Benthic Species in Southern Regions While Greatest Economic Impacts Occur in Northern Regions.” *Ecological Modelling* 383 (10): 106–17.

Marshall, Kristin N., Isaac C. Kaplan, Emma E. Hodgson, Albert Hermann, D. Shallin Busch, Paul McElhany, Timothy E. Essington, Chris J. Harvey, and Elizabeth A. Fulton. 2017. “Risks of Ocean Acidification in the California Current Food Web and Fisheries: Ecosystem Model Projections.” *Global Change Biology* 23 (4): 1525–39.

Bauer B., Meier H. E. M., Casini M., Hoff A., Margoński P., Orio A., *et al.* (2018). Reducing eutrophication increases spatial extent of communities supporting commercial fisheries: a model case study. ICES J. Mar. Sci. doi:10.1093/icesjms/fsy003.

Stäbler M, Kempf A, Mackinson S, Poos JJ, Garcia C, Temming A (2016) Combining efforts to make maximum sustainable yields and good environmental status match in a food-web model of the southern North Sea. Ecol Model 331:17–30, DOI:10.1016/j.ecolmodel.2016.01.020

Cormon X, Kempf A, Vermard Y, Vinther M, Marchal P (2016) Emergence of a new predator in the North Sea: evaluation of potential trophic impacts focused on hake, saithe, and Norway pout. ICES J Mar Sci 73(5):1370–1381, DOI:10.1093/icesjms/fsw050

5.2 ToR B. Update of key-runs (standardized model runs updated with recent data, producing agreed output and agreed upon by WGSAM participants) of multi-species and ecosystem models for different ICES regions (Baltic Sea EwE, LeMans Framework proposed for use in Irish Sea)

WGSAM has performed two key runs: the Ewe Baltic and The North Sea SMS model key run. In addition, the model framework of LeMans was reviewed.

North Sea SMS model key-run (2017)

In 2017, a Key Run of the North Sea Stochastic Multispecies Model (SMS) was presented and reviewed in detail by four WGSAM experts, and approved by the group following implementation of changes agreed in plenary at the meeting and verified by a subset of experts post-meeting. The SMS model was produced using data from the period 1974–2016. This included updates to the input data and some modification to the structure of the model. These are described in detail in the stock annex:

http://www.ices.dk/sites/pub/Publication%20Reports/Stock%20Annexes/2017/StockAnnex_ICES_NS_SMS_Configuration.pdf

The main results of the 2017 key-run can be found in:

(http://ices.dk/community/Documents/Expert%20Groups/WGSAM/NS-keyRun.zip).

Model code, input and output can be found at the ICES expert group Github:

(https://github.com/ices-eg/wg_WGSAM).

Key run summary sheet – North Sea SMS

Area	North Sea
Model name	SMS
Type of model	Age-length structured statistical estimation model
Run year	2017
Predatory species	Assessed species: Cod, haddock, saithe, whiting, mackerel 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
Prey species	Cod, haddock, herring, Norway pout, southern North Sea sandeel, northern North Sea sandeel, sprat, whiting,
Time range	1974–2016
Time step	Quarterly
Area structure	North Sea
Stomach data	Fish species: 1981, 1985, 1986, 1987, 1991, 2005, 2013 Grey seals: 1985, 2002

	Harbour porpoise: Decadal 1985, 1995, 2005
Purpose of key run	Making historic data on natural mortality available and multispecies dynamic
Model changes since last key run	All time-series updated. Mackerel included as a modelled stock. Proportion of the stock within the North Sea given as input and used for estimating M2. Daily food ration of changed for the main fish species. Bias correction of diet composition of harbour porpoise and the main predatory fish.
Input and output available at	Sharepoint/data/North_Sea_key_run and from the ICES expert group Github (https://github.com/ices-eg/wg_WGSAM).
Further details in	Report of the Working Group on Multispecies Assessment Methods 2017

Use of results by ICES stock assessments.

One of the main results from the North Sea SMS is the estimates of predation mortality (M2) which together with an estimate of residual narural mortality (M1) are used as natural mortality (M=M1+M2) by the ICES single stock assessment. M values from SMS are used by the following stocks in the North Sea area: Cod, haddock, whiting, herring, sprat and sandeel. As an example, the herring M2 values are shown in figure SMS-1.

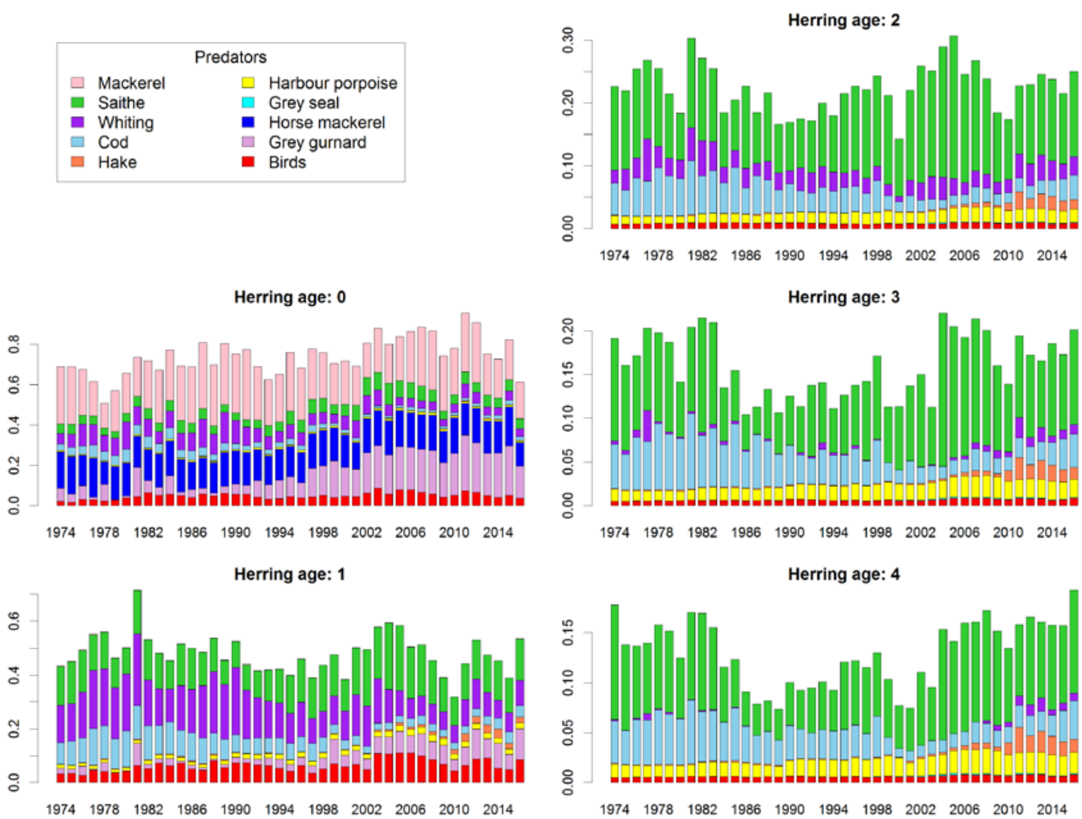


Figure SMS-1 Annual predation mortality (M2) by age of herring inflicted by predator species.

Identified areas of priority research

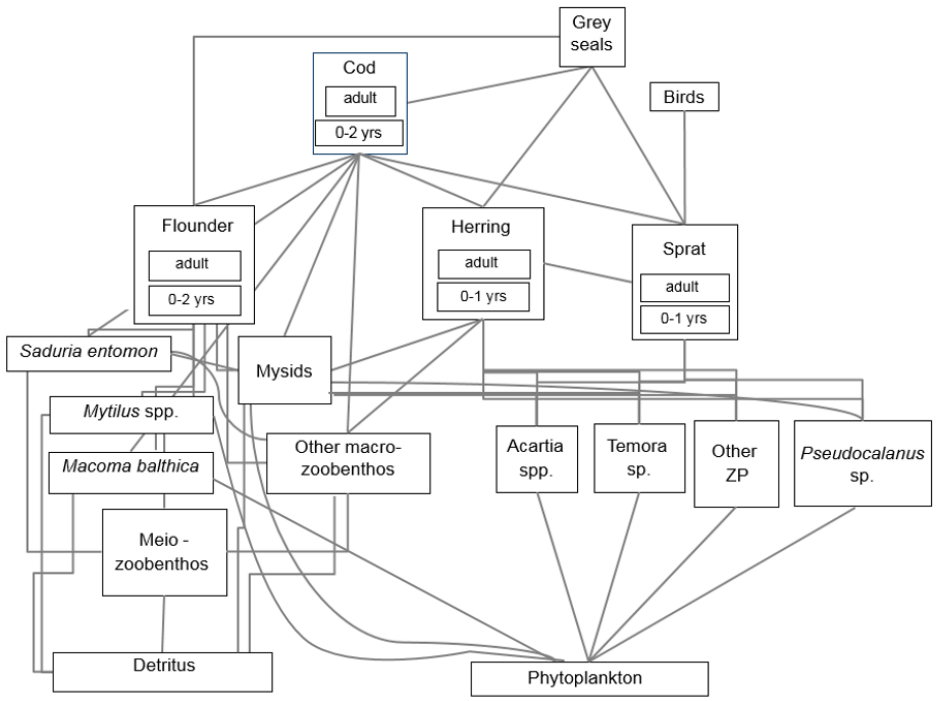
WGSAM considers that the following topics should be priority areas of study prior to the next North Sea key run:

WGSAM 2017 priority areas of study prior to the next North Sea key run:	**WGSAM 2018 action**
Estimating the proportion of hake, mackerel and horse mackerel stocks present in the North Sea and their **distribution in northern and southern areas** for a better estimation of M2 for the two sandeel stocks.	Under the FISHPI2 project, there is an effort to produce new data on distribution of all fish stocks by identifying the necessary sampling and produce sampling manuals. It is likely that the expansion of the swept area survey for mackerel in 2018 may provide new data to support estimates of distribution of mackerel. ACTION: FISHPI2 manual to be commented by WGSAM chair
Estimating distributions of seabirds in southern and northern North Sea.	This requires contributions from e.g. WGBIRD based on the seabirds at sea database. Most likely to provide useful results if WGSAM sends a scientist who can participate in the next meeting in October 2019. ACTION: WGSAM will send a request to JWGBIRD
Reviewing the **method used to estimate grey gurnard and starry ray abundance** to identify the reference period and sizes to which the average biomass estimates apply. Consider if the SMS model by it likelihood statistics can estimate a likely mean biomass over a given period.	This requires modelling work. ACTION: To be investigated before WGSAM 2019
Update the number of seabirds, grey seals and harbour porpoise with the most recent information.	Seabirds requires contributions from e.g. JWGBIRD based on the seabirds at sea database and nesting populations. Most likely to provide useful results if we are able to find a person who can participate in the next meeting (they have just met). Grey seals and harbour porpoise have historically been provided by Sophie Smout, St. Andrews. ACTION: contact taken to St. Andrews and request send to JWGBIRD
Consider using annual harbour porpoise data.	It is possible that annual harbour porpoise data may be better at reflecting differences in prey abundance than the current decadal averages. ACTION: contact taken to St. Andrews

Update the diet and consumption data for grey seal with the most recent data.	These data are available at St. Andrews. Recent studies using genetics indicate that the otoliths substantially underrepresent smaller fish specimens. ACTION: contact taken to St. Andrews
Assigning prey to length groups for the 2013 mackerel stomach data.	This requires modelling work. Data was sent last year but not included in the key run. ACTION: To be investigated before WGSAM 2019
Establishing quarterly catch histories for the all predator species (cod, whiting, haddock, saithe, mackerel) as initiated with data from InterCatch.	Data are available in RDB format in the North Sea RCG. ACTION: A request to the RCG to be drafted and sent.
Investigate changes to modelling performance when including **overwintering mortality of sandeel** (M1, possible condition or weight at age dependent).	This should be investigated in single species assessments before being included in multi-species versions.
Estimating the abundance in the sea of small fish.	This task is complicated. It could be possible to explore methods back-calculating growth
Investigate the most appropriate **species and size selection** of different predators.	This requires modelling work and a reliable estimate of the abundance of small fish. Hence, it awaits a solution to the above task.
Work towards obtaining **new stomach data.**	Under the FISHPI2 project, there is an effort to produce new data for relevant predator stocks by identifying the necessary sampling and produce sampling manuals. The work requires an estimate of the number of stomachs required by species and a manual for scaling the content of individual stomachs to North Sea scale food consumption. ACTION: FISHPI2 manual to be commented by WGSAM chair
Including localised stomach sampling in models.	This will be possible if the model begins to include specified spatial aspects.

Baltic Sea EwE model Keyrun (2016)

The WG presented a first Ecopath with Ecosim key run for the Central Baltic Sea in 2016. The run is a development of the model described by Tomczak *et al.* (2012) and Niiranen *et al.* (2013), but parameterised according to post-regime shift conditions. The model includes one phytoplankton group, four meso-zooplankton groups (*Pseudocalanus* spp., *Acartia* spp., *Temora* spp., and 'other mesozooplankton' which consists of other copepods and cladocerans), one epi-benthic group represented by Mysids, five benthos groups (*Saduria entomon, Macoma balthica, Mytilus* spp., meiobenthos and 'other macrozoobenthos'), four fish species (eastern Baltic cod *Gadus morhua*, central Baltic herring *Clupea harengus*, Baltic Sea sprat *Sprattus sprattus* and flounder *Plathychtys flesus*), grey seal *Halichoerus grypus* and offshore fish-feeding birds (razorbill *Alca torda*, common guillemot *Uria aalge* and black guillemot *Cepphus grille)*.



Model structure – Baltic Sea EwE

Three types of forcing are applied in the model: fishing effort (cod and flounder), fishing mortalities (sprat and herring) and abiotic forcing (cod, herring, sprat, phytoplankton, zoonplankton). The model was evaluated with PreBal diagnostics which are an established and standardized method to test the quality of mass-balance models within the 'EwE community'. Input data and model structure are described in detail in the stock annex http://www.ices.dk/sites/pub/Publication%20Reports/Expert%20Group%20Report/SSGEPI/2016/01%20WGSAM%20-%20Report%20of%20the%20Working%20Group%20on%20Multispecies%20Assessment%20Methods.pdf#search=EwE%202016.

The main results of the 2016 key-run can be found in http://ices.dk/community/Documents/Expert%20Groups/WGSAM/EwE_BS_KeyRun%202016%20Outputs.zip.

So far, the model contribution to advice and management has been limited. Given the nature of the EwE models, its potential contribution should be expected on evaluation of long-term consequences of alternative managemet actions which may extend beyond fisheries management (i.e., management of nutrient loads).

The model has been preliminarily compared to other multispecies models for the central Baltic Sea and used to inform the prototype of a decision support tool for ecosystem based fisheries management in the central Baltic Sea within the context of a research project. Further, more extensive comparisons with key runs of other more quantitative models (i.e., SMS, Gadget) should be expected in the near future.

Key run summary sheet – Baltic Sea EwE

Area	Baltic Sea
Modelling approach	Ecopath with Ecosim
Type of model	Foodweb compartment
Run year	2016
Species/Groups	22 functional groups
Time range	2004–2013
Time-step	Yearly (internal multistanza calculations monthly)
Area structure	Model covers approximately Baltic Proper ICES Subdivisions 25–29, excl. 28–1. A spatial extension of the model is under development.
Stomach data	From the EU tender "Study on stomach content of fish to support the assessment of good environmental status of marine foodwebs and the prediction of MSY after stock restoration"
Purpose of key run	Description of changes in the Baltic Sea foodweb
Model changes since last key run	First key run

5.3 ToR C. Consider methods to assess the skill of multispecies models intended for operational advice

Several approaches focussing on different aspects of skill assessments for multi species and ecosystem models were presented during the meetings of WGSAM 2016, 2017 and 2018. The approaches ranged from testing the predictive power of diet selection and consumption sub-models, how multi species models compare to single species approaches up to performance testing of whole models in an MSE approach. A simulation study was conducted to investigate how diet data quality, availability and the method used to fit diet data impact the performance of multispecies assessment models. The analyses presented also included sensitivity tests of models to input data as well as model structure. New applications of global sensitivity analyses of model results to parameters were developed. In some case studies model predictions were challenged with observed hindcast time-series or retrospective patterns in hindcasts and forecasts were analysed. One of the main messages was that the quality of hindcasts does not allow for conclusions on the predictive power of the model. Details of the presentations and main conclusions can be found in the Annex for TOR C of the WGSAM 2017 and 2018 report.

A cooperation with ICES WGIPEM on skill assessments of complex models was agreed and a summary presentation on skill assessments and keyruns carried out by members of WGSAM was presented at the WGIPEM meeting in 2018. Collaboration to write a best practice paper and general work together with WGIPEM will continue in the next years to come up with guidelines for skill assessments that can be applied in ICES.

Keyruns are a core activity of WGSAM and they refer to a model parameterization and output that is accepted as a standard by ICES WGSAM, and thus serves as a quality assured source for scientific input to ICES advice. Although skill assessments are carried out, the process can be improved once more complete guidelines for skill assessments (especially also for the forecast part) of complex models become available. The importance of detailed documentation of input, model settings and diagnostics next to skill assessments has also to be highlighted. WGSAM uses Github, an extra stock annex on the ICES website, standardized main output (tables and figures) and puts effort into the direct communication with e.g., assessment working groups.

As final conclusion, progress has been achieved in providing an overview of latest developments in skill assessment methods. However, more work is needed to derive final best practice guidelines from the various approaches available. Therefore, it is recommended to continue this ToR in the next three years and strengthen the collaboration with WGIPEM and similar groups that work on similar topics.

Relevant papers

Spence, M. A., Blackwell, P. G. and Blanchard, J. L. 2016. Parameter uncertainty of a dynamic multispecies size spectrum model. Can. J. Fish. Aquat. Sci. 73: 589–597. dx.doi.org/10.1139/cjfas-2015–0022

Bauer B., Horbowy J., Rahikainen M., Kulatska N., Müller-Karulis B., Tomczak M.T., Bartolino V. Sources of structural uncertainty and its impacts on simulated fisheries management scenarios in the Baltic Sea. PlosOne, in review.

Kulatska N, Neuenfeldt S, Beier U, Elvarsson BÞ, Wennhage H, Stefansson G, Bartolino V. Understanding ontogenetic and temporal variability of Eastern Baltic cod diet using a multispecies model and stomach data. Fish. Res., in review.

Lehuta S, Girardin R, Mahévas S, Travers-Trolet M, Vermard Y. Reconciling complex system models and fisheries advice: Practical examples and leads. Aquat Living Resour. avr 2016;29(2):8.

5.4 ToR D. Investigate the performance of multi-model ensemble in comparison to single model approach

Spence *et al.* (2018) developed a general framework for combining ecosystem models. The model is probabilistic and therefore quantifies the uncertainty in estimates. This was presented in 2016, 2017 and an application demonstrating the effectiveness of three management strategies was presented in 2018.

An interactive multispecies model, T-ONS, was designed for stakeholders to use. In it a multispecies Schaefer model, fitted to SMS outputs, is converted to a Jacobian matrix, which is a linear approximation to the response surface at status quo effort. This was presented in 2017.

Five Climate-Enhanced (CE) models (stock, multispecies, ecosystem, fleet and human community) will be used together to evaluate potential responses to projected climate change in the eastern Bering Sea. Alternative management strategies will be evaluated under different climate projections. This is ongoing work and was presented in 2017.

A multi-model approach was used to understand the role of Pacific sardine in the California Current food web. Although not a true ensemble approach, this work highlighted the influence of structural assumptions, such as taxonomic resolution, age structure, and density dependence, on the predictions of the models. This was presented in 2017.

A simple Multi-Model Inference (MMI) approach was applied to two US Atlantic herring stock assessment models to "field test" the approach for use during a benchmark assessment process. This was presented in 2018.

A generic algorithm was used to get multiple plausible parameter values for the surplus production model Hydra, a length structured multispecies model with explicit recruitment and predation. This was presented in 2018.

Relevant Papers

Spence, M.A., Blanchard , J.L., Rossberg, A.G. *et al.* (2018) A general framework for combining ecosystem models. Fish and Fisheries. 2018; 00:1–12. https://doi.org/10.1111/faf.12310.

Pope J.G., Bartolino V., Kulatska N., Bauer B., Horbowy J., Ribeiro J.P.C., Sturludottir E., Thorpe R. Comparing the steady state results of a range of multispecies models between and across geographical areas by the use of the Jacobian matrix of yield on fishing mortality rate. Fish. Res., in press.

5.5 ToR E. Test performance and sensitivity of ecosystem indicators

Over the course of the three-year period, WGSAM has considered two new ecosystem indicators, a food-web evenness and Species-Area-Relationship (SAR) indices, as well as discussed ecosystem thresholds, indicator performance, and time-series trend detection.

The food-web evenness index accounts for the loss of energy and biomass towards higher trophic levels. It is used to measure the ecosystem state in relation to when specific species within a trophic level or specific trophic level are disproportionally abundant. While it has been applied to several systems, the application of the food-web evenness index to the Baltic Sea Ecopath model (Tomczak *et al.*, 2012) was presented to the working group. This indicator is expected to be suitable for analyzing simulated management scenarios by informing managers about the state of the ecosystem with systems likely being more unstable with disrupted function when biomass at certain trophic levels strongly decline (Carpenter *et al.*, 1985; Prugh *et al.*, 2009) or when one or a few species dominate a trophic level (Atkinson *et al.*, 2014).

The SAR Index is a measure of the rate at which species accumulate with increasing area. It was tested along with the Large Fish Index (LFI) and Mean Trophic Level (MTL) on the Swedish west coast including Skagerrak and Kattegat to assess marine ecosystem status and outcomes of management actions such as Marine Protected Areas (MPAs) and trawl limit regulations. The SAR approach provides novel insights into the way we perceive and understand changes at the community and ecosystem level in relation to biodiversity loss, fishery and governance (Novaglio *et al.* 2016 submitted). The SAR slope may be viewed as a novel indicator to be further considered and explore within the EU MSFD descriptors D1 and D4.

In addition to the two new ecosystem indicators, the working group discussed ecosystem thresholds, indicator performance, and time-series trend detection. A challenge in working with ecosystem indicators is that both ecosystem drivers and ecosystem responses can be complex and multidimensional. Using several empirical datasets, gradient forests were used to evaluate whether common drivers acting together might lead to common multivariate responses in ecosystems, and whether ecosystem thresholds could be identified using multiple indicators responding to multiple drivers. Then, dynamic factor analysis was used to further characterize and analyze ecosystem trends, and generalized additive models were used to illustrate another method to identify threshold responses to individual drivers such as fisheries landings across the four ecosystems. Thresholds identified with these multiple methods were reasonably robust within an ecosystem. While outcomes of these empirical methods are somewhat dependent on time-series length and quality, the overall methods are promising for integrating multiple drivers and evaluating cumulative effects of both human activities and environmental pressures.

Understanding the limitations with respect to the development of management advice using indicators of varying lengths, as well as our ability to comment on "recent" trends, is important. The ability to identify statistically significant trends from time-series of varying length and autocorrelation regimes was assessed using Monte-Carlo simulations. Time-series with varying degrees of trend and autocorrelation as well as length were simulated. Three statistical approaches (Mann-Kendall, Mann-Kendall with pre-whitening, and Generalized Least Squares) were then tested for their ability to correctly identify trends in the data. Results indicate that our ability to identify recent trends in the data is limited, and statistical approaches were biased by even moderate amounts of autocorrelation at small samples sizes (N < 30). In addition, a recent paper shows that PCA is not robust to autocorrelations, so other methods should be explored instead. This will be explored by WKINTRA. Methods will be similar to the above, applying different analytical tools to simulated time-series with known qualities to examine the strengths and weaknesses of different methods.

Relevant papers

Fay, G., Large, S. I., Link, J. S., and Gamble, R. J. 2013. Testing systemic fishing responses with ecosystem indicators. Ecological Modelling, 265: 45–55.

Fay, G., Link, J. S., Large, S. I., and Gamble, R. J. 2015. Management performance of ecological indicators in the Georges Bank finfish fishery. ICES Journal of Marine Science: Journal du Conseil, 72: 1285–1296.

Hardison, S., Perretti, C., DePiper, G., and Beet, A. In prep. A simulation study of trend detection methods for IEA.

Large, S. I., Fay, G., Friedland, K. D., and Link, J. S. 2013. Defining trends and thresholds in responses of ecological indicators to fishing and environmental pressures. ICES Journal of Marine Science: Journal du Conseil, 70: 755–767.

Large, S. I., Fay, G., Friedland, K. D., and Link, J. S. 2015. Quantifying Patterns of Change in Marine Ecosystem Response to Multiple Pressures: e0119922. PLoS One, 10. http://search.proquest.com/docview/1664222461/abstract/120C1207C54548BEPQ/1 (Accessed 14 October 2016)

Novaglio, C., Svedäng, H., Sköld, M., Belgrano. A. In prep. Species-Area Relationship (SAR) and Marine Protected Areas (MPAs): linking fish communities status and marine conservation measures.

Tam , J.C., Link, J.S., Large, S.I., Andrews, K., Friedland, K.D., Gove, J., Hazen, E., Holsman, K., Karnauskas, M., Samhouri, J.F., Shuford, R., Tomilieri, N., and Zador, S. 2017. Comparing apples to oranges: common trends and thresholds in anthropogenic and environmental pressures across multiple marine ecosystems. Frontiers in Marine Science, 4: 282. https://doi.org/10.3389/fmars.2017.00282

5.6 ToR F. Metanalysis of impact of top predators on fish stocks in ICES waters

Although marine mammals are major drivers of marine ecosystems, diet and consumptions studies are often lacking. Reviews of diet and consumption data on both sides of the Atlantic indicates that marine mammal consumption data is sporadic, erratic, and rather ad-hoc. There are few long-term time-series, many populations are represented by a small number of one-off investigations. WGSAM notes that this lack of a consistent data underpinning makes modelling marine mammal consumption and ecosystem impacts problematic.

WGSAM 2017 investigated the importance of top-predators in ICES waters for selected case studies where consistent data are available.

The impact of harbour porpoise and grey seal on commercial fish species in the North Sea was investigated in the ToR B North Sea keyrun with a specific focus on correcting for bias in stomach content (underestimation of small fish). Mammals were found to be the primary natural mortality source of e.g. cod in recent years and their importance has increased considerable since 1974 (the key run start year).

The preliminary results from the GADGET model for cetacean fishery interaction in the Iberia peninsula suggested that the multispecies model improves the quality of the fit compared with a hake single species model; the biomass of hake consumed by cetaceans are on the same scale as historical hake catches. The multispecies model could be used to provide advice for hake considering impact on hake and cetaceans.

For further information regarding the results from ToR F refer to Annex 8: ToR F; and ToR B North Sea keyrun (WGSAM 2017).

Relevant papers

Trijoulet, V., Holmes, S. J., and Cook, R. M. 2018. Grey seal predation mortality on three depleted stocks in the West of Scotland: What are the implications for stock assessments? Canadian Journal of Fisheries and Aquatic Sciences, 75: 723–732.

Trijoulet, V., Dobby, H., Holmes, S. J., and Cook, R. M. 2018. Bioeconomic modelling of grey seal predation impacts on the West of Scotland demersal fisheries. ICES Journal of Marine Science, 75: 1374–1382.

Silber, G.K., M.D. Lettrich, P.O. Thomas, J.D. Baker, M. Baumgartner, E.A. Becker, P. Boveng, D.M. Dick, J. Fletcher, J. Forcada, K.A. Forney, R.B. Griffis, J.A. Hare, A.J. Hobday, **D. Howell**, K.L. Laidre, N. Mantua, L. Quakenbush, J.A. Santora, K.M. Stafford, P. Spencer, C. Stock, W. Sydeman, K. Van Houtan, and R.S. Waples. 2017. Projecting marine mammal distribution in a changing climate. Frontiers in Marine Science 4:413. doi:10.3389/fmars.2017.00413

5.7 ToR G. Explore the consequence of multispecies, mixed fisheries interactions and environmental factors in practical multispecies advice for fisheries management (MSY related and other biological reference points)

WGSAM has a long-standing experience in discussing fisheries management in a multi species context. During this three years period: the Nash equilibrium for the Baltic fish community has been revisited; the WGSAM was made aware of a new project on "Eco-system Based FMSY Values in Fisheries Management", which outputs have been tracked during these years; the EBFM approaches in the USA based on ceilings for total system removals or species complexes; as part of the development of the roadmap for an EAF in the NAFO area has, the advancements in the multispecies approach have been presented, the performance of the Nash equilibrium in comparison to the ICES Fmsy approach has been compared in terms of biological risk versus economic revenues.

Ecosystem Fmsy project

The Ecosystem Fmsy project has been conducted to attempt to find practical steps for improving the current Fmsy/Ftarget fishing levels by including more ecosystem realism. Attempting to produce full multispecies advice has not, in general, been successful in entering tactical management, but a more realistic method of computing single species Fmsy target values may be easier to implement in management. The project is now in the writing-up phase, and held a short final symposium in Copenhagen at which two WGSAM members held presentations. Production (and Ecopath) models were run for a large number of ICES stocks in order to include density dependence and hence implicitly ecosystem considerations, into target fishing levels. These production models produced alternate candidate "multispecies" Fmsy values, which form a first step towards improving the realism in precautionary evaluated ICES HCRs.

WGSAM supports this effort to improve the current management targets in this way. Although the WG considers that it is clearly not precautionary to uncritically adopt a set of production model Fmsy values for a range of stocks, the group does consider that using these values to identify a subset of stocks to conduct further MSE-style evaluations with a view to potentially revising existing target reference points represents a viable approach to improving ICES management.

The approach for ecosystem-based management in the Northeast United States.

The approach for and EBM in two of the Northeast USA Regional Management Councils (RMC) was presented during WGSAM 2016. In the New England RMC the North-east Fisheries Science Center (NEFMC) is developing a fisheries ecosystem plan (FEP), with the strategy of setting an overall system level cap on removals with individual species protection. In first place the Ecosystem Production Units (EPUs) are defined. Then an overall fisheries catch cap is set using a simple energy flow model to determine ecosystem production potential or other modelling approaches. Fishery functional groups defined as those species that are caught together by specified fleet sectors and that play similar roles in the ecosystem with respect to energy transfer. Accordingly, the concept developed by the NEFMC in New England encapsulates information on technological interactions as well as trophic guild structure and feeding interactions. The final step involves specifying catch levels for Individual species without exceeding the functional group cap (catch ceiling) and none of the individual species are exploited at unsustainable levels. In the Mid-Atlantic RFC the NEFMC approach retains the current single species Fishery Management Plans (FMPs), but incorporates relevant climate, habitat, predator, and other interactions as possible. The framework for integrating ecosystem approaches into current fishery management was built on aspects of the Integrated Ecosystem Assessment approach. Recently, a risk assessment was completed for the Mid-Atlantic which will be used to further specify integrated modelling analyses and possibly multispecies management strategy evaluation.

New England Herring MSE

A management strategy evaluation (MSE) to test harvest control rules that consider Atlantic herring's role as forage in the Northeast US shelf ecosystem was presented. This may be the first MSE in the US Fishery Management Council process to hold a public stakeholder workshop to generate objectives and performance measures and to identify key sources of uncertainty to be considered in the analysis. The New England Fishery Management Council selected a herring harvest control rule based on this MSE in September 2018.

Evaluating an ecosystem-based fishery management procedure for Georges Bank using ceilings on system removals.

Closed loop simulation was conducted to test a proposed Ecosystem-Based Fishery Management (EBFM) strategy for Georges Bank in the Northeast U.S. A ceiling on total system removals was implemented in the MSE runs with indicator-based harvest control rules to evaluate combinations of management actions, which may be effective for managing multiple species at once in an ecosystem context. It was found that the ceiling level on total system removals explains most of the variability in performance metrics at the whole ecosystem, aggregate species groups, and single species level. In addition, the implementation of indicator-based harvest control rules also explained a large portion of performance variability when ceilings were set to higher values.

Ecosystem Based Fishery Management in New England, USA

A proposed EBFM procedure was described, with components illustrated using model simulations. Six steps were outlined:

- Need of identifying spatial management units
- Establish specific management objectives and exploitation reference points directed at stock complexes rather than individual species.
- Establish biomass thresholds (floors) below which the complex as a whole cannot fall (Option 1) or below which no species within the complex can fall (Option 2).
- Devise an Ecosystem-based Harvest Control Rule based on steps 2 and 3 designed to minimize the risk of overfishing for a range of exploitation rates at the stock complex level.
- Simulate the performance of a set of scenarios constructed under the EBMP using a suit of metrics including biomass, landings, revenue, probability of

breaching a threshold biomass level, maintaining robust size structure of the populations (large fish index), and the stability of the landings.

- Identify and reconcile trade-offs.

The results showed that low levels of exploitation rate (0.15–0.20) had the best overall performance in terms of biomass, yield, and risk of falling into the depleted status. It was concluded that the major trade-offs involve catch, revenue, and species-complex or species status.

The common fisheries policy and Nash equilibrium MS–MSY

The Nash equilibrium (NE) among stock harvest rates proposed during the WGSAM meeting in 2015 as multispecies MSY reference points (ICES 2015) were re-estimated for cod, herring and sprat in the eastern Baltic Sea using a Multi-Species Interaction Stochastic Operative Model (MSI-SOM) (Norrström *et al.*, 2016). The MSI-SOM was run for two scenarios: the FMSY and BMSY as target. Targeting Bmsy produced more appealing results within the MS-MSY framework, although it would imply a dcrease in catches for some stocks compared to the current SS-MSYs.

Nash equilibrium to understand productivity in a multispecies system during environmental change – The Baltic Sea as a case study

The effects of the climate change scenarios of salinity and temperature in the Baltic from the report by Meier *et al.* (2012) on the productivity of a multispecies system were explored using the Nash equilibrium estimate of FMSYs, due to its capacity to synthesize the importance of ecological interactions between the species in a multispecies system and the environmental drivers (Norrström *et al.* 2016). The environmental changes were shown to have little effect on the productivity of the sprat stock (<10% increase) but stronger effect on the cod (~50% reduction) and especially the herring stock (~80% reduction).

The effects of density dependent clupeid growth on Nash equilibrium reference points in the Baltic Sea

In the 2017 meeting it was shown that fitting of the clupeid growth functions in the MSI-SOM model to data is improved with intra- and interdependent density dependence. The inclusion of this density dependent effect had a minor effects on the Nash equilibrium reference points.

Multispecies Management Strategy Evaluation

It would obvious be valuable to be able to test out potential Harvest Control Rules (single or multispecies) against a multispecies Operating Model. To date this has been hindered by a lack of available tools. Rather than attempt to build ad-hoc MSE tools around a given multispecies Operating Model, the REDUS project in Norway has aimed to link existing simulating models (including multispecies ones) with existing MSE tools. Combining state-of-the-art tools in this way minimizes the potential for introducing errors, and allows individual components to be updated independently of each other and hence to keep the overall tool up to date. Computer code has been developed to link Gadget to the A4A Management Strategy Evaluation tool, and the Flemish Cap multispecies model has been used as a test case for this code. Work is ongoing to extend this by linking the model to FLBEIA. Both FLBEIA and A4A are FLR-based, and are both used for research and management within Europe.

Multispecies approach in NAFO. Management Strategy Evaluation in Flemish Cap as a case study

As part of the EU SC05 project "Multispecies Assessment for NAFO fisheries" an updated version of the multispecies model GadCap (Pérez-Rodríguez *et al.* 2017) was presented. The results of the model showed the joint effect of predation and fishing has produced the strong changes in biomass observed for cod, redfish and shrimp. A multispecies MSE framework integrating the multispecies model GadCap as operating model within an a4a-MSE framework was also presented. Single species and multispecies oriented HCRs were design and tested using this multispecies MSE framework. The results showed the influence that variable fishing strategies on predators (cod and redfish) would have on the prey stocks (shrimp and redfish). It was concluded that single species oriented management strategies were not precautionary for cod and shrimp. The results also suggested that it is not possible having the 3 sps above Blim, and that disregarding one stock (shrimp or another stock) may allow finding precautionary multispecies reference points for the others.

Multispecies management strategy evaluation in the North sea. A comparison of PGY ranges, ices 2012 assessments, and the nash equilibrium

A management strategy evaluation (MSE) being conducted using a length-structured multispecies and mixed fisheries model of the North Sea fish community (LeMans, 2015, 2016, 2017) was presented. The Fmsy values from the ICES 2012 stock assessment, 21-stock Nash equilibrium, one based on single species assessments, and ones based upon the top, middle, and bottom of the ICES "pretty good" yield ranges were evaluated using a variety of Harvest Control Rules (HCR), with outcomes being assessed in terms of average risk of stock depletion and gross revenue (price x catch). It was concluded that the Nash is generally better than ICES 2012, and PGY lower is safer, whilst PGY upper is worse.

Balanced Harvesting

Balanced Harvesting (BH) is a proposed fishery management regime in which fishing is targeted across the widest feasible range of ecosystem components in proportion to their productivity (by species and by size). In contrast to traditional management, BH would target smaller species and younger fish (which have higher productivity) over the older individuals of larger species as at present. It is likely that BH could lead to increased yield in biomass from an ecosystem with reduced impact on the ecosystem structure. However, this increased yield may come from less commercially valuable sizes and species, and by spreading harvesting to a wider range of target sizes and species BH may raise fishing costs. It is therefore not clear that BH is desirable, and more detailed simulation modelling is therefore required.

WGSAM has been involved in the BH debate through several publications (Howell *et al.* 2016, and a review of BH currently under review) discussing the impacts and implications of BH, and through a project to simulate the impacts of possible BH on the Barents Sea through Atlantis modelling (paper in prep).

SYMBIOSES oil-risk project

WGSAM has collaborated with the SYMBIOSES project to develop an integrated oil-risk assessment tool for the Lofotens in Norway. This tool combines detailed modelling of oceanography, oil transport, ecotoxicology, and IBM larval drift and survival in the Lofotens with a multispecies fish model in the Barents Sea. This work is now published (Carroll *et al*. 2017) with cod as an example species, a further funding application in under consideration to expand this other species spawning in the area.

Relevant papers

Gaichas, S. K., DePiper, G. S., Seagraves, R. J., Muffley, B. W., Sabo, M., Colburn, L. L., and Loftus, A. L. 2018. Implementing Ecosystem Approaches to Fishery Management: Risk Assessment in the US Mid-Atlantic. Frontiers in Marine Science, 5. https://www.frontiersin.org/articles/10.3389/fmars.2018.00442/abstract (Accessed 20 November 2018).

Gaichas, S. K., Fogarty, M., Fay, G., Gamble, R., Lucey, S., and Smith, L. 2017. Combining stock, multispecies, and ecosystem level fishery objectives within an operational management procedure: simulations to start the conversation. ICES Journal of Marine Science, 74: 552–565.

Norrström N., Casini M., Holmgren N.M.A., 2016, Nash equilibrium can resolve conflicting maximum sustainable yields in multi-species fisheries management, ICES Journal of Marine Science, doi:10.1093/icesjms/fsw148

Pérez-Rodríguez, A., Howell, D., Casas, M., Saborido-Rey, F. Ávila-de Melo, A. 2016. 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, doi:10.1139/cjfas-2016–0111

Thorpe, R.B., P.J. Dolder, S. Reeves, P. Robinson, and S. Jennings (2016) Assessing fishery and ecological consequences of alternate management options for multispecies fisheries, ICES Journal of Marine Science, 2016, DOI 10.1093/icesjms/fsw028

Thorpe, R.B., S. Jennings, P.J. Dolder (2017) Risks and benefits of catching pretty good yield in multispecies mixed fisheries, ICES Journal of Marine Science, DOI 10.1093/icesjms/fsx062

Carroll, J., Vikebø, F., Howell, D., Broch O.J., Nepstad, R., Starrlight, A., Skeie, G.M., Bast, R, and Juselius, J. 2017. Assessing impacts of simulated oil spills on the Northeast Arctic cod fishery. Marine Pollution Bulletin 126, pp_63–73, doi.org/10.1016/j.marpolbul.2017.10.069

Gullestad P., Howell D., Stenevik E.K., Sandberg P. and G. Bakke. Management and rebuilding of herring and cod in the Northeast Atlantic. 2018. In Garcia, S.M. and Ye, Y., eds. 2018. Rebuilding of marine fisheries. Part 2. Case studies. FAO Fisheries and Aquaculture Technical Paper No. 630/2. Rome, FAO. 234 pp.

Howell, D., Hansen, C., Bogstad, B., and Skern-Mauritzen M. 2016. Balanced harvesting in a variable and uncertain world – a case study from the Barents Sea. ICES Journal of Marine Science. doi: 10.1093/icesjms/fsw034

Tallman, R., Janjua M. Y., Howell D., Ayles, B., Carmicheal, T., Bernreuther, M., Ferguson S., and Treble, M. 2016. Pan.Arctic Fisheries and their Assessment. In "Fisheries and Aquaculture in the Modern World", ed. Heimo, M. ISBN 978–953–51–2687–4, Print ISBN 978–953–51–2686–7. doi: 10.5772/64745

Deroba, J. J., Gaichas, S. K., Lee, M.-Y., Feeney, R. G., Boelke, D. V., and Irwin, B. J. 2018. The dream and the reality: meeting decision-making time frames while incorporating ecosystem and economic models into management strategy evaluation. Canadian Journal of Fisheries and Aquatic Sciences. http://www.nrcresearchpress.com/doi/10.1139/cjfas-2018–0128 (Accessed 20 July 2018).

Kempf A, Mumford J, Levontin P, Leach A, Hoff A, Hamon KG, Bartelings H, Vinther M, Stäbler M, Poos JJ, Smout S, Frost H, Burg S van den, Ulrich C, Rindorf A (2016) The MSY concept in a multi-objective fisheries environment - lessons from the North Sea. Mar Policy 69:146–158, DOI:10.1016/j.marpol.2016.04.012

6 Cooperation

Cooperation with other WGs:

- Working Group on Integrative, Physical-biological and Ecosystem Modelling (WGIPEM)
- Working Group on Mixed Fisheries Advice (WGMIXFISH)
- Workshop on integrated trend analyses in support to integrated ecosystem assessment (WKINTRA)
- Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK)
- Herring Assessment Working Group for the Area South of 62° N (HAWG)
- Baltic Fisheries Assessment Working Group (WGBFAS)

Cooperation with Advisory structures

WGSAM keyruns provided natural mortality estimates for various stocks in the North Sea and Baltic

7 Summary of Working Group self-evaluation and conclusions

WGSAM continues providing a forum for developing the approaches, methods and tools to support ICES in providing integrated, ecosystem-based advice.

This report summarizes the achievements of the group in the last 3 years, the details of which are documented in three reports covering the meetings in 2016–2018.

The self-evaluation form highlights that more should be done to better integrate the groups’ advice-relevant outputs in to the ICES advisory system. This is particularly important to address now given the proposed ToRs for 2019–2021 (particularly ToR e) and their relationship to the commitments made in the CFP and the MSFD to implementation of the ecosystem approach to management.

The group requests a continuation of the WG for a new 3-year term.

Annex 1: List of participants

Name	Institute	Email
Daniel Howell	Institute of Marine Research P.O. Box 1870 Nordnes Norway	danielh@imr.no
Alexander Kempf (co-chair)	Thuenen Institute of Sea Fisheries Palmaille 9 22767 Hamburg Germany	alexander.kempf@thuenen.de
Andrea Belgrano	Swedish University of Agricultural Sciences, Department of Aquatic Resources, Institute of Marine Research, Turistgatan 5, SE-453 30 Lysekil, Sweden and Swedish Institute for the Marine Environment (SIME), Box 260, SE-405 30 Göteborg, Sweden	andrea.belgrano@slu.se
Robert Thorpe (correspondence)	Centre for Environment, Fisheries and Aquaculture Science (CEFAS) Pakefield Road NR33 0HT Lowestoft Suffolk United Kingdom	robert.thorpe@cefas.co.uk
Morten Vinther	DTU-Aqua Charlottenlund Casttle 2920 Charlottenlund Denmark	mv@aqua.dtu.dk
Valerio Bartolino	Swedish University of Agricultural Sciences, Department of Aquatic Resources, Institute of Marine Research, Turistgatan 5, SE-453 30 Lysekil, Sweden	valerio.bartolino@slu.se
Alfonso Perez Rodriguez	Wageningen Marine Research Research Scientist	alfonso.perezrodriguez@wur.nl
Sarah Gaichas (co-chair)	NOAA Northeast Fisheries Science Center, Woods Hole USA	Sarah.Gaichas@noaa.gov

Michael Spence	Centre for Environment, Fisheries and Aquaculture Science (CEFAS) Pakefield Road NR33 0HT Lowestoft Suffolk United Kingdom	michael.spence@cefas.co.uk
Ching Villanueva (remote)	French Institut of Research for the Exploitation of the Sea (Ifremer), Pointe du Diable, 29280 Plouzané	Ching.Villanueva@ifremer.fr
Vanessa Trijoulet	Integrated Statistics under contract to NOAA Fisheries Northeast Fisheries Science Center Population Dynamics Branch 166 Water St. Woods Hole, MA 02543, USA	vanessa.trijoulet@nooa.gov
Anna Rindorf	DTU-Aqua Charlottenlund Casttle 2920 Charlottenlund Denmark	ar@aqua.dtu.dk
Sigrid Lehuta	French Institut of Research for the Exploitation of the Sea (Ifremer), rue de l'ile d'Yeu, 44300 Nantes	slehuta@ifremer.fr
Sean Lucey	Dept. of Commerce/NOAA/NMFS Northeast Fisheries Science Center 166 Water Street, Woods Hole, MA 02543	sean.lucey@noaa.gov
Hayley Bannister	Centre for Environment, Fisheries and Aquaculture Science (CEFAS) Pakefield Road NR33 0HT Lowestoft Suffolk United Kingdom	
Klaas Sys	Institute for Agricultural and Fisheries Research (ILVO) Belgium	klaas.sys@ilvo.vlaanderen.be

Annex 2: Recommendations

RECOMMENDATION	ADRESSED TO
1. Development in quarterly population numbers of feeding seabirds of the following species in the North Sea in the period from 1974 to 2018, preferably divided into proportion feeding north and south of 56.5 oN: Fulmar, gannet, Greater blackbacked gull, Guillemot, Herring gull, Kittiwake, Puffin and Razorbill.	JWGBIRD

Annex 3: WGSAM resolution 2019–2021

The **Working Group on Multispecies Assessment Methods** (WGSAM), chaired by Alexander Kempf, Germany, and Sarah Gaichas, USA, will work on ToRs and generate deliverables as listed in the Table below.

	Meeting dates	Venue	Reporting details	Comments (change in Chair, etc.)
Year 2019	14–18 October	Rome, Italy	Interim report by 1 December	
Year 2020			Interim report by DATE	
Year 2021			Final report by DATE	

ToR descriptors

ToR	Description	Background	Science Plan codes	Duration	Expected Deliverables
a	Review further progress and deliver key updates on multispecies modelling and ecosystem data analysis contributing to modelling 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.
b	Update of key-runs (standardized model runs updated with recent data) of multispecies and eco-system models for different ICES regions	The key runs provide information on natural mortality for inclusion in various single species assessments	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	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	Establish methods 2019, apply 2020–2021	Manuscript for methods, report on success of methods for different examples.
d	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, as applied for example management questions	5.1; 6.1; 6.3	3 years	Report on methods for comparing models and for constructing model ensembles.

e	Management Strategy Evaluation (MSE) methods and applications for mutispecies and ecosystem advice, including evaluating management procedures and estimating biological reference points	Adapting existing multispecies/ecosystem models for MSE (operating models, assessment models), visualizing tradeoffs and uncertainty for managers and stakeholders	5.3; 6.1; 6.3	3 years	Review of MSE modelling approaches. Review of visualization methods. Review of applications throughought the ICES area with lessons learned.

Summary of the Work Plan

Year 1	All ToRs, Key run Baltic, multiple models
Year 2	All ToRs, Key Run North Sea SMS (maybe others)
Year 3	All ToRs, Key Run US Northeast Shelf, multiple models

Supporting information

Priority	The current activities of this Group will lead ICES into issues related to the ecosystem effects of fisheries, especially with regard to the application of the MSY Approach. The activities will provide information (e.g., natural mortality estimates, performance of indicators) and tools (e.g., multi-model ensembles, keyrun models) valuable for the implementation of an integrated advice in several North Atlantic ecosystems. Consequently, these activities are considered to have a very 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	Approx 20. Expertise in ecosystem, modelling and fish stock assessment from across the whole ICES region.
Secretariat facilities	None.
Financial	No financial implications.
Linkages to ACOM and groups under ACOM	ACOM, most assessment Expert Groups
Linkages to other committees or groups	WGMIXFISH, WGDIM, WGBIFS, IBTSWG, WGECO, WGINOSE, WGIAB, WGNARS, WGIPEM.
Linkages to other organizations	None

Annex 4: WGSAM self-evaluation 2016–2018

1) Working Group name: Working Group on Multispecies Assessment Methods (WGSAM)

2) Year of appointment: 2016

3) Current Chairs: Alexander Kempf (Germany) and Sarah Gaichas (USA)

4) Venues, dates and number of participants per meeting:

Reykjavik, Iceland, 10–14 October 2016 (17 participants)

San Sebastian, Spain, 16–20 October 2017 (21 participants)

Paris, France, 15–19 October 2018 (16 participants)

WG Evaluation

5) If applicable, please indicate the research priorities (and sub priorities) of the Science Plan to which the WG make a significant contribution.

WGSAM members conduct and publish research on multispecies and multi-fisheries interaction and thus their work is closely linked with the two ICES science goals and their specific activities, specifically in order of priority:

Goal 2

Understand the relationship between human activities and marine ecosystems, estimate pressures and impacts, and develop science-based, sustainable pathways

- developing integrated ecosystem assessment methodologies and approaches that allow the use of both qualitative and quantitative data, and which can be used to address both specific advisory questions and broader ecosystem issues;
- providing tools and methods for assessing the relationships between marine ecosystems, their biological resources, and the provision of services (particularly food security) to society, including socio- economic aspects;

Goal 1

Develop an integrated, interdisciplinary understanding of the structure, dynamics, and the resilience and response of marine ecosystems to change

- investigating the structure, functioning, dynamics, and interconnectedness of marine ecosystems, their different biotic components, and the abiotic environment at different spatial scales;

6) In bullet form, list the main outcomes and achievements of the WG since their last evaluation. Outcomes including publications, advisory products, modelling outputs, methodological developments, etc. *

- WGSAM executed key runs for the Baltic Sea (Ecopath with Ecosim in 2016) and the North Sea (SMS in 2017). One of the main results from the North Sea SMS is the estimates of predation mortality (M2) which together with an estimate of residual natural mortality (M1) are used as natural mortality (M=M1+M2) by the ICES single stock assessment. M values from SMS are used by the following stocks in the North Sea area: Cod, haddock, whiting, herring, sprat and sandeel.
- Numerous papers and reports have been published by WGSAM members relevant to the ToRs between 2016–2018 (see under chapter 5 ToR outcomes above).

7) Has the WG contributed to Advisory needs? If so, please list when, to whom, and what was the essence of the advice.

- WGSAM executed key runs for the Baltic Sea (Ecopath with Ecosim in 2016) and the North Sea (SMS in 2017). One of the main results from the North Sea SMS is the estimates of predation mortality (M2) which together with an estimate of residual natural mortality (M1) are used as natural mortality (M=M1+M2) by the ICES single stock assessment. M values from SMS are used by the following stocks in the North Sea area: Cod, haddock, whiting, herring, sprat and sandeel.

8) Please list any specific outreach activities of the WG outside the ICES network (unless listed in question 6). For example, EC projects directly emanating from the WG discussions, representation of the WG in meetings of outside organizations, contributions to other agencies' activities.

- Collaborations in EU H2020 project proposals (various) catalysed through the network opportunities afforded by WGSAM, but no specific project arising from WG discussions.
- Reviews of ecosystem modelling programmes in the USA. (New England EBFM simulation framework review)
- Work presented and refined at WGSAM has been presented in a variety of fora (e.g. ICES ASC, American Fisheries Society, Internal Research Institute presentations).

9) Please indicate what difficulties, if any, have been encountered in achieving the workplan.

We foresee that the possibility of a growing demand for evaluation of new models and their applications as Key Runs (some planned with multiple models) could become a heavy draw which impacts the broader work of WGSAM. To ensure financing by member countries it may be beneficial to bring WGSAM (at least partly) under the ACOM umbrella.

Future plans

10) Does the group think that a continuation of the WG beyond its current term is required? (If yes, please list the reasons)

Yes.

Reasons:

- M2 values from key runs are an essential component of stock advice in the North Sea and the Baltic.
- Article 9 of the CFP specifically Article 9,3b on multiannual plans which states "Multiannual plans shall cover: in the case of mixed fisheries or where the dynamics of stocks relate to one another, fisheries exploiting several stocks in a relevant geographical area, taking into account knowledge about the interactions between fish stocks, fisheries and marine ecosystems".
- The MSFD, particularly GES descriptor 4 Food Webs, requires information on how biological and fishery interactions affect the functioning of food webs and the consequences for ecosystem and its capability for provisioning services.
- Policy in other ICES jurisdictions (including the USA and Norway) also commits to adoption of ecosystem approach and the development of relevant methodologies to implement this.
- The work of WGSAM is intimately linked to the ICES Strategic Plan goal 3 on sustainable use: "Scientific information is the foundation of ICES advice and this advice must meet the needs of decision-makers. ICES will continue to deliver evidence- based scientific advice on environmental is sues and fishery management. ICES is committed to transition, where appropriate, from single-species fisheries advice to advice in a mixed fishery, multispecies, and ecosystem context. ICES is developing regional integrated advice based on ecosystem assessments including indicators for assessing ecosystem status, and for the management of human activities.

11) If you are not requesting an extension, does the group consider that a new WG is required to further develop the science previously addressed by the existing WG. N/A

(*If you answered YES to question 10 or 11, it is expected that a new Category 2 draft resolution will be submitted through the relevant SSG Chair or Secretariat*.)

12) What additional expertise would improve the ability of the new (or in case of renewal, existing) WG to fulfil its ToR?

- Continuation of the wide range of multispecies experience and expertise is critical. In addition, cooperation with other WGs (WGMIXFISH, WGECO, WGIPEM) should be extended.

13) Which conclusions/or knowledge acquired of the WG do you think should be used in the Advisory process, if not already used? (please be specific)

- In addition to the current use of key runs to provide M2 values, these models (preferably as ensembles where multiple models occur for a region) should be used to provide advice on the possible ecosystem effects of MSY policy and of changes to specific targets/ management reference points. This could include both effects on stocks, fisheries and ecosystem indicators, and could extend to management strategy evaluation using multispecies and ecosystem models.

Annex 5: ToR A. Review further progress and deliver key updates in multispecies and ecosystem modelling throughout the ICES region

The review of progress of multispecies models in ICES Ecoregions given below is not intended to be comprehensive and exhaustive. It reflects the knowledge available to the participants at the 2018 meeting and input from WGSAM who were not able to attend in person.

There was no participation from Russia or Canada at this year's meeting, and consequently no update on modelling from the regions.

Ecoregion A: Greenland and Iceland Seas

There is no progress to report on multispecies modelling in the Ecoregion this year.

Ecoregion B: Barents Sea

Work has been developed in IMR Bergen under the REDUS project to link a Gadget multispecies Operating Model to the A4A MSE framework. This allows for harvest control rules, both single and multispecies, to be evaluated against a multispecies reality. This has been implemented using the multispecies model for the Flemish cap, with results described under ToR g. This method of linking state-of-the art systems leverages the strength of both. Compared to producing ad hoc tools for each situation, linking existing tools in this manner also makes maintaining up-to-date tools easier.

The Atlantis model for the Norwegian and Barents Seas has been further developed, including a study to the investigate potential implications that would arise if Balanced Harvesting were to be applied to the Barents Sea.

See updates under ToR f (impact of top predators), and ToR g (multispecies and mixed fishery advice).

Ecoregion C: Faroes

There is no progress to report on multispecies modelling in the Ecoregion this year.

Ecoregion D: Norwegian Sea

Progress here is presented in conjunction with work in the Barents Sea under ecoregion B.

Ecoregion E: Celtic Seas

Avoiding the curse of circularity: building a multi-species model from the ground up·

- A multi-species size-based model of the Celtic Sea with 17 species was fitted, with quantifiable uncertainty, to landings and survey data without the use of single-species assessments.
- Fishing rates were fitted for each species for 24 years by taking advantage of the model structure.

- Emergent spawning stock biomass estimates similar assessment estimates for cod, whiting and hake, but different for haddock.
- The model could be useful to validate candidate assessment models, give advice for data limited stocks or to make long-term forecasts.

Spence presented the multi-species size-based model mizer (Blanchard *et al.* 2014) fitted to 17 species in the Celtic Sea. Biological parameters were taken from survey data and fishing rates, maximum recruitment and background resource were fitted to IBTS and landings data. By treating the fishing rates as “tuning” parameters they lost meaning but he emergent SSB estimates were comparable to the single species SSB (Figure 1).

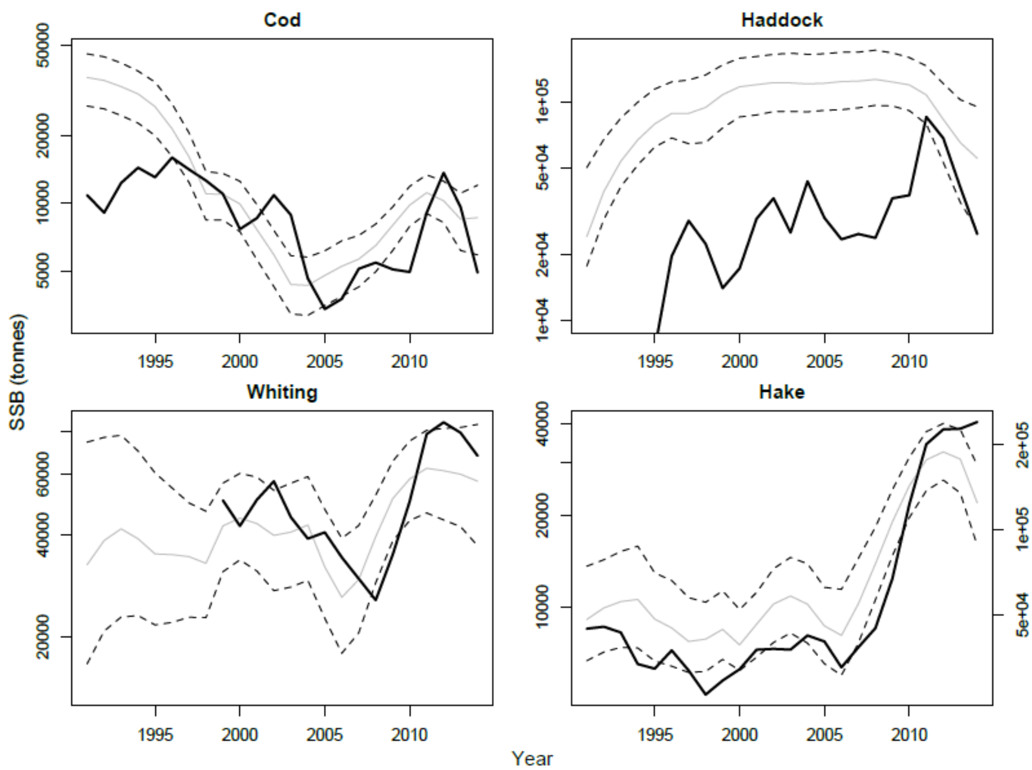


Figure 1. The median modelled SSB, the 10 and 90 percentiles and SSB estimates from single-species ICES assessments for cod, haddock, whiting and hake. The hake assessment covers more area than the model does and therefore is plotted on a different scale.

For more details email Michael Spence: michael.spence@cefas.co.uk

Blanchard, J.L., Andersen, K.H., Scott, F., Hintzen, N.T., Piet, G., and Jennings, S. 2014. Evaluating targets and trade-offs among fisheries and conservation objectives using a multispecies size spectrum model. J. Appl. Ecol. 51(3):612–622. doi:10.1111/1365–2664.12238.

Spence, M. A., Blackwell, P. G. and Blanchard, J. L. 2016. Parameter uncertainty of a dynamic multispecies size spectrum model. Can. J. Fish. Aquat. Sci. 73: 589–597. dx.doi.org/10.1139/cjfas-2015–0022

Ecoregion F: North Sea

Multispecies and mixed fisheries management strategy evaluation in the North Sea

Robert Thorpe gave a presentation on a management strategy evaluation (MSE) being conducted using a length-structured multispecies and mixed fisheries model of the North Sea fish community. Five candidates for a community MSY (CMSY), the 21-stock Nash equilibrium, one based on single species assessments, and ones based upon the top, middle, and bottom of the ICES “pretty good” yield ranges were evaluated using a variety of Harvest Control Rules (HCR), with outcomes being assessed in terms of average risk of stock depletion and gross revenue (price x catch). The MSE was carried out with an ensemble of 63 models with stochastic recruitment, repeated 100 times for each scenario. In the absence of an HCR, we find that the lower PGY ranges are the safest option and the Nash equilibrium the highest yielding, with the other options being sub-optimal. Application of an HCR cuts risk and reward, the former more than the latter such that the HCR is useful. The impact of the HCR depends on its functional form and the point at which yield is reduced (MSY Btrigger). We find that the optimum choice for CMSY depends on societal views of acceptable risk, with no clearly optimum solution. However, the upper part of the PGY ranges is never a good choice.

Previous work with LeMans (Thorpe 2015, 2016, 2017) has focused on constant harvesting strategies – fishing at the same mortality regardless of stock status to evaluate the long-term impacts of the strategy. But in practice, fishing would not take place in this way, because it would be reduced if the stock status is poor. This is often done via a harvest control rule (HCR), a pre-agreed management procedure which determines how the fishing mortality target varies with stock status. In this study we ask what is the best way of achieving multispecies MSY (assuming this is defined as achieving the maximum possible yield for an acceptable risk using HCRs within a management strategy evaluation framework. Thus, we take account of model parameter uncertainty, management target uncertainty, and fleet management uncertainty as well as management implementation uncertainty, evaluating outcomes in terms of risk and reward. A schematic of the experiment design is shown in Figure 1.

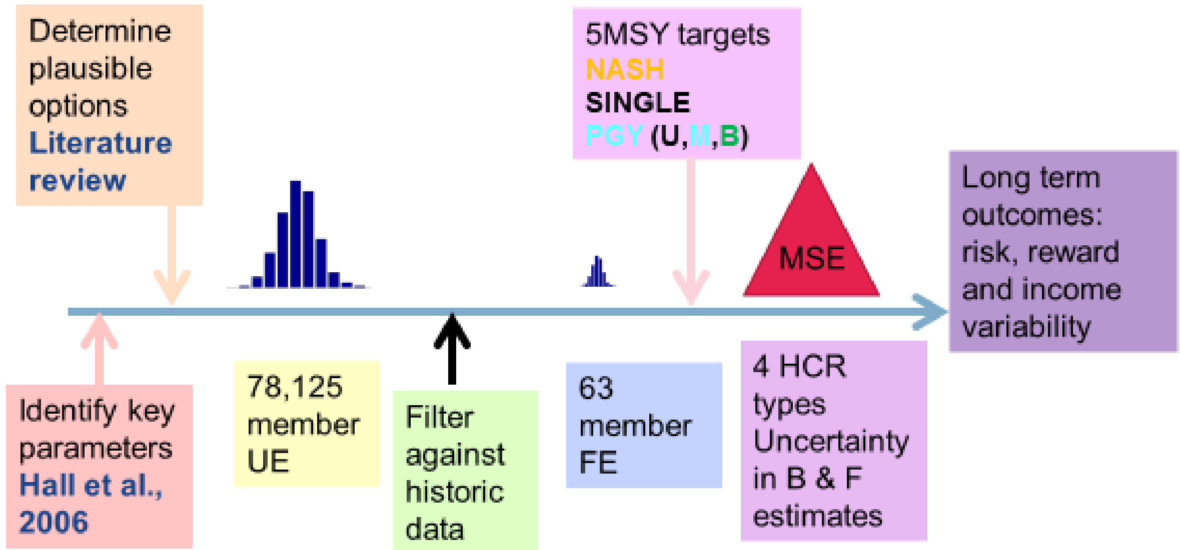


Figure 1. Schematic of the experimental design for the management strategy evaluation.

We use the same 63 member ensemble as in Thorpe *et al.* (2017), with each ensemble being evaluated 100 times to take account of stochastic variation in recruitment. 5 candidates for a community MSY were considered, one based on 2012 single species assessments (Thorpe *et al.* 2015), one based on the Nash equilibrium (Thorpe *et al.* 2017) and three based on the bottom, middle, and top of the "pretty good yield" ranges as defined by ICES. Of the 21 stocks, 7 (cod, haddock, whiting, sole, plaice, herring, and saithe) have published ranges as in Table 1.

Table 1. ICES estimates of "pretty good yield" for 7 North Sea stocks.

Stock	F-PGY upper	F-PGY central	F-PGY lower
Herring	0.39	0.33	0.24
Sole	0.37	0.20	0.113
Whiting	0.15	0.14	0.14
Plaice	0.30	0.21	0.146
Haddock	0.194	0.194	0.167
Cod	0.46	0.31	0.198
Saithe	0.49	0.36	0.21

The other 14 PGY ranges were generated using the following assumptions. 1) The average F across all stocks was maximised, 2) no fleet can have more than three times the effort of another, when they are both expressed as effort relative to the average effort between 1990 and 2010, and 3) there are no discards, so the fishery is limited as soon as the first choke limit is reached. Figure 2 shows the Fs that result from these assumptions.

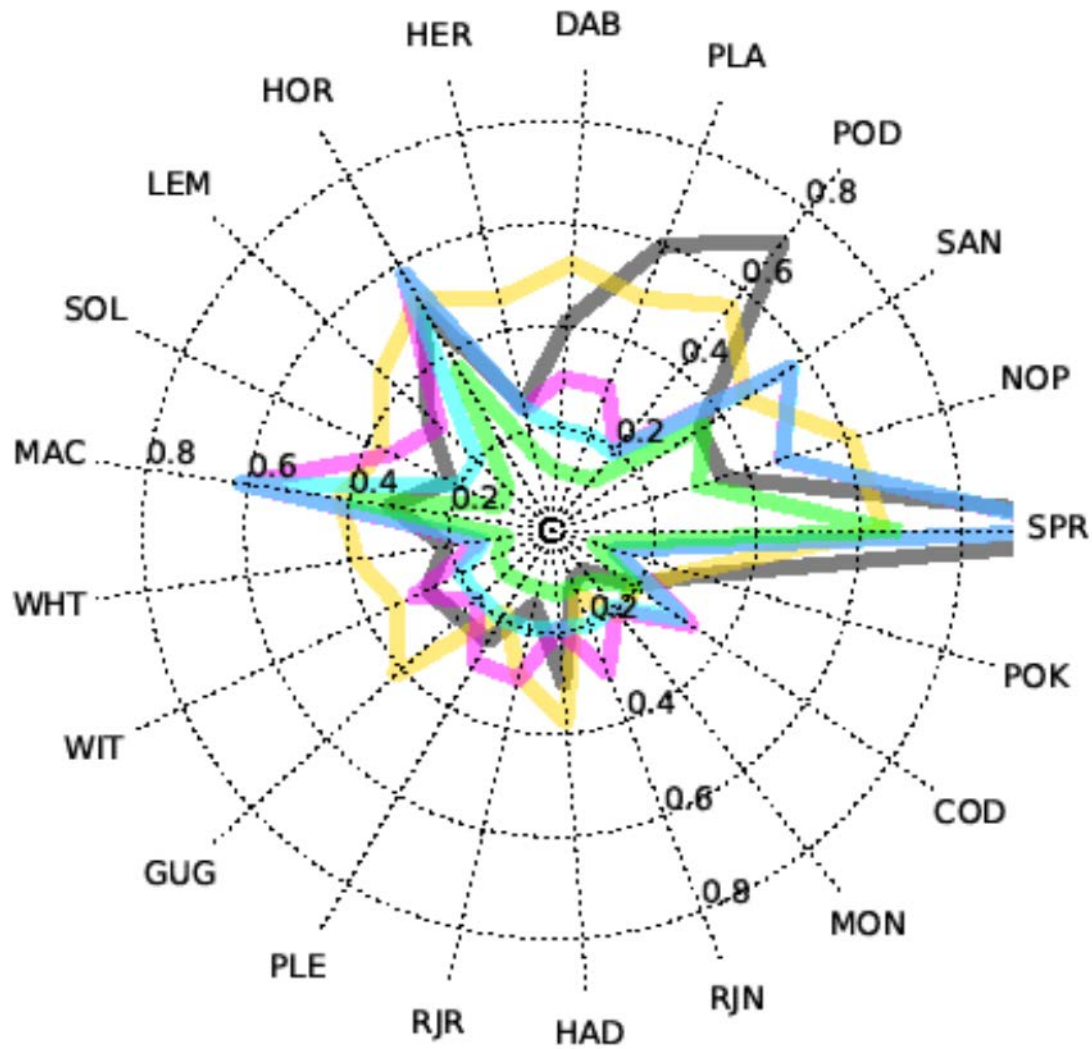
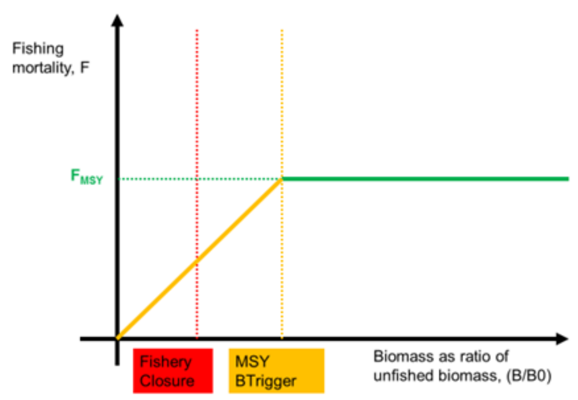
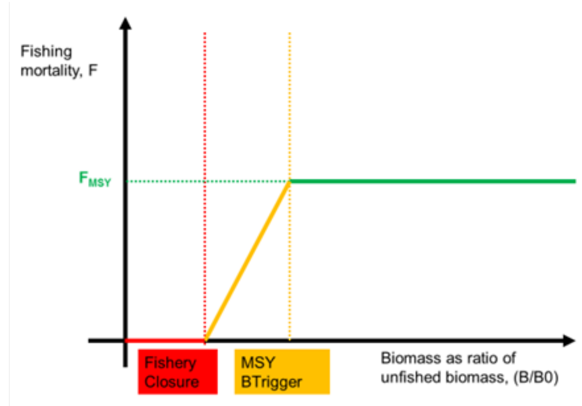


Figure 2. Radar plot of community MSY for a) estimates based on single species assessments (black), b) 21-stock stochastic Nash equilibrium (gold), c) upper PGY ranges (magenta), d) mid PGY ranges (cyan), and e) lower PGY ranges (green).

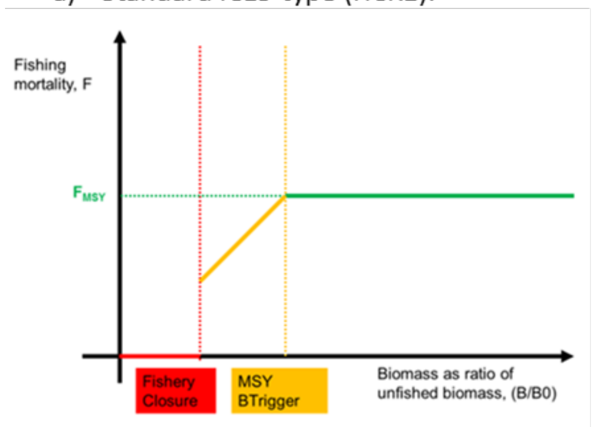
Our management strategy evaluation considered 4 types of harvest control rules, which are shown in Figure 3.



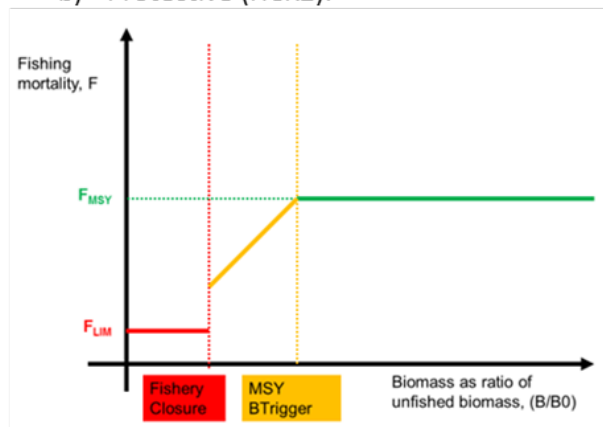
a) Standard ICES-type (HCR1).



b) Protective (HCR2).



c) Precautionary (HCR3).



d) Realistic (HCR4).

Figure 3. Schematic of the 4 types of harvest control rules used in the management strategy evaluation.

Within the MSE, the ensemble model acts as the operating model. Stock status is assessed by taking the ensemble mean biomasses adjusted by a log-normal error term of given size from 0 to 50%. The harvest control rule is then used to generate a target F for the stock, which is implemented with lognormal uncertainty of given size from 0 to 30%. Stock status is assessed annually, after which the newly ascertained F is applied for the next year.

For the reference case with constant F and no HCR, results are shown in Figure 4.

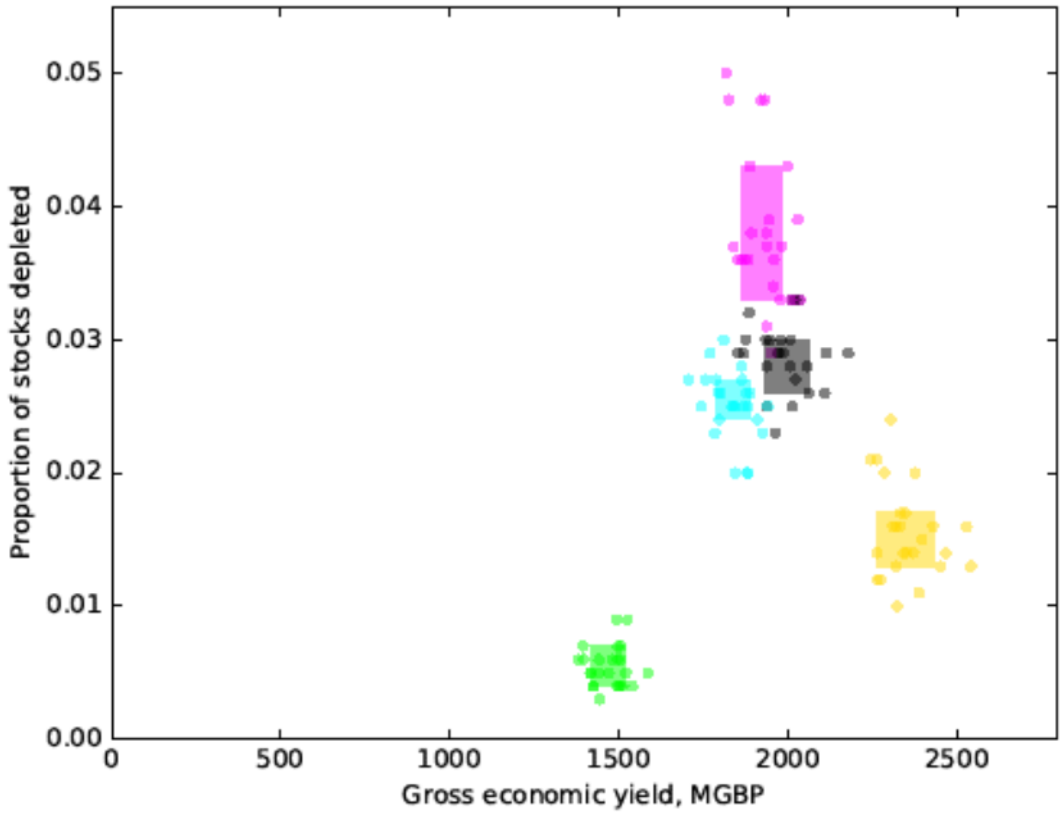


Figure 4. Risk – reward outcomes for constant F strategies. Black = single species, Gold = Nash, GreEn = L-PGY, Cyan = M-PGY, Magenta = U-PGY.

We find that the Nash equilibrium gives the highest yield, and L-PGY is the safest. A choice between these would depend on societal risk appetite. The other solutions are sub-optimal. Application of an HCR reduces both risk and yield, the former more than the latter. The nature of the reduction depends upon the form of the HCR, the choice of MSY Btrigger, and the level at which a stock is considered depleted.

Overall we find the following:-

a) The best outcome depends upon societal views of risk and reward – there is no choice of CMSY that is clearly optimal, although some can be dismissed as sub-optimal.
b) The upper PGY ranges are never a good choice (consistent with Thorpe *et al.* 2017).
c) The annual operation of an HCR reduces both risk and yield. Yield reductions are more modest than those of risk, making the HCR a valuable management tool, independent of its form (amongst those considered here), definition of risk, or definition of MSY Btrigger.

References:

S.J. Hall, J.S. Collie, D.E. Duplisea, S. Jennings, M. Bravington, and J. Link (2006) A length-based multispecies model for evaluating community responses to fishing. Canadian Journal of Fisheries and Aquatic Sciences, 63, 1344–1359.

R.B. Thorpe, W.J.F. Le Quesne, F. Luxford, J.S. Collie and S. Jennings (2015) Evaluation and management implications of uncertainty in a multispecies size-structured model of population and community responses to fishing, Methods in Ecology and Evolution, 6, 49–58. doi: 10.1111/2041–210X.12292

R.B. Thorpe, P.J. Dolder, S. Reeves, P. Robinson, and S. Jennings (2016) Assessing fishery and ecological consequences of alternate management options for multispecies fisheries, ICES Journal of Marine Science, 2016, DOI 10.1093/icesjms/fsw028

R.B. Thorpe, S. Jennings, P.J. Dolder (2017) Risks and benefits of catching pretty good yield in multispecies mixed fisheries, ICES Journal of Marine Science, DOI 10.1093/icesjms/fsx062

Ecospace for the southern part of the North Sea

A southern North Sea Ecopath with Ecosim (EwE) has been finalized at the Thünen Institute of Sea Fisheries (TI-SF) by Staebler *et al.* to a fitted and calibrated stage. Publications in peer reviewed journals are available. Based on this model Puets *et al.* developed an Ecospace model at TI-SF. The model is used to explore spatial management strategies. Novel is that time dynamic maps for habitat suitability can now be included to take into account changes in distribution and habitat suitability over time. Different model runs were carried out to test whether the frequency of updating these maps has an influence on model performance. The scenario that just uses one map from the start year as input (current default) performed worse compared to model runs updating the maps every five years, each year or even each season (if available like for maps based on IBTS data). Therefore, it is important to take into account changes in habitat suitability over time.

As first scenarios, it was tested whether the closure of 1) operational windparks, 2) operational and planned windparks and 3) windparks and MPAs (see figure 1) for demersal fisheries has a positive impact on fish stocks. Preliminary results indicate that the effects are non-linear and especially for roundfish the impact overall was even negative (figure 2). The reason is that effort displacement and food web effects outside the closed areas overrule the positive effects inside the closed areas. Therefore, impact assessments for spatial management require models that take into account the most important processes and models that can only predict linear first order effects are not sufficient.

In the near future, the sensitivity of model results to parameter input (e.g., dispersal rates) will be tested. Hot spots of biodiversity as well as sensitivity will be investigated to give guidance on optimised spatial management. Climate change scenarios will be tested to see how spatial management measures would need to be adapted under climate change.

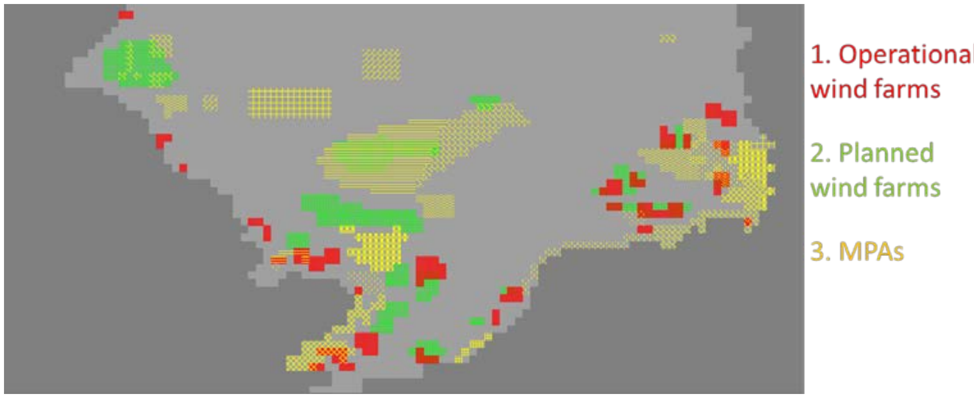


Figure 1. Operational and planned wind farms as well as MPAs included in the Ecospace sceanrios.

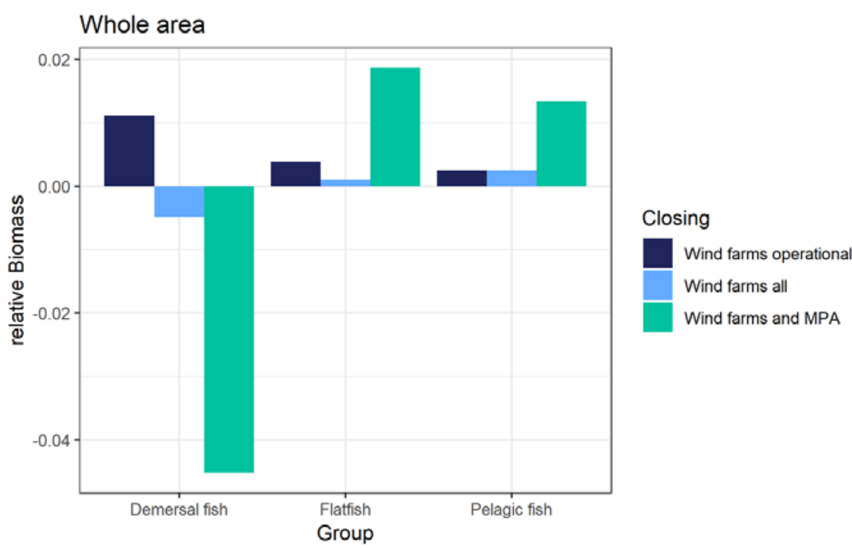


Figure 2. Relative biomass change (%) compared to no closure by fish group and scenario.

Ecoregion G: South European Atlantic Shelf

ISIS-Fish updates

- ISIS-Fish is a spatial simulation focused on technical interactions
- It is currently applied to five case studies to answer various questions including assessing marine managed areas network, management plans and management measures such as the landing obligation, MSE loops, and exploring climate change effects and integrated ecosystem assessment opportunities.
- The model is increasingly used in collaboration with stakeholders. Supporting material (animated movie, interactive interfaces…) was presented that helped communicate model assumptions, functioning and results to the sector.

The activity around the ISIS-Fish model was presented to the group. ISIS-Fish is a spatial multi-species model that does not account for trophic interactions but only technical interactions. It is currently applied to 5 cases study: The strait of Sicily, the hake fishery in the Gulf of Lion, the Bay of Biscay pelagic and demersal fisheries, and the Eastern English Channel.

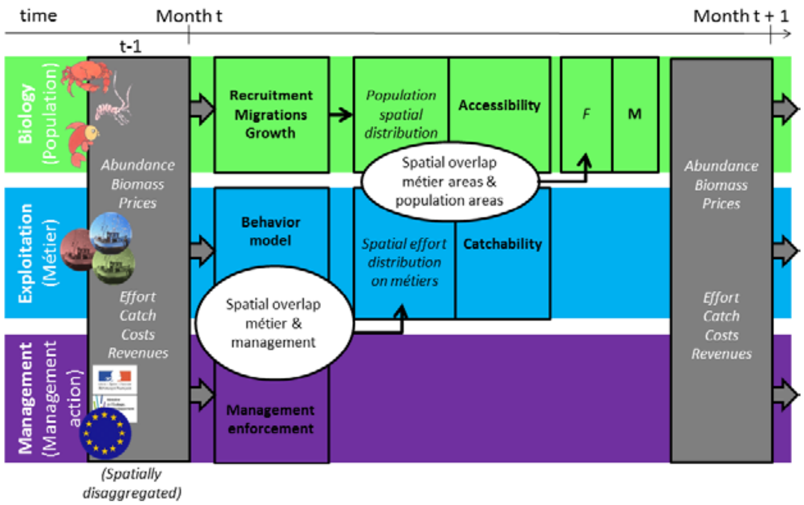


Figure 1. Conceptual diagram of ISIS-Fish model functioning highlighting the exchanges of information between the management, fishing activity and population modules within a time step.

The questions that motivated the modelling work were diverse. In the strait of Sicily, the performances of a network of Marine Managed Areas will be tested in the context of the MANTIS project. In the Gulf of Lion the model aims at investigating the spatial structure of the hake population which experiences a high fishing pressure. The uncertainties on stock spatial distribution and reproductive potential prevent the assessment of the management measures designed for the fishery. The ISIS-Fish application of the demersal fishery of the Bay of Biscay models the technical interactions between sole, nephrops and hake. It has been recently used as the operating model in an MSE loops using the Stock synthesis model for hake to evidence the uncertainties in hake assessment dur to hake spatial structure and fleet dynamics (A. Vigier, phD thesis 2018). The pelagic model of the Bay of Biscay is focused on the impact of climate change on pelagic species and the propagation of the effects to the fleets (EU CERES project). A fleet dynamic model was developed to account for pelagic fleets long-term changes in strategies. The biological module will be fed with outputs from a DEB-IBM model, itself forced with ERSEM results of IPCC scenarios in order to simulate the impact of the climate at the population scale and on the fleets long-term dynamics. The ISIS-Fish model of the mixed demersal fishery of the Eastern English Channel models the dynamics of 10 species and 17 fleets. It has been recently used in the DiscardLess project to evaluate the impact of the landing obligation. In the context of this project the model was presented and discussed with the stakeholders together with the Osmose and Atlantis models of the area. Communication supports (movie, leaflets, interactive interfaces) were developed to ease the discussion with fishers. A theoretical fleet dynamics model was created to predict fishers' response to changes in fish availability and market opportunities (Lehuta *et al.* 2015). It assumes that behaviour depends for a part α on fishing habits and responds to catch opportunities for the rest. In order to evaluate the impact of the assumptions regarding fleet behaviour three values of α were tested. The model was used to assess scenarios of discard allowance, landing obligation and de minimise exemptions. The simulations evidence gains in revenues at the fishery scale in the long-term with the landing obligation and even more with the exemptions. In the short term however, losses of about 10% per year are expected. The results are actually contrasted between fleets some being beneficiary and others in deficit. The opportunity for reallocation of effort also varied across fleets. The results are being presented to stakeholders in late 2018.

Lehuta S, Vermard Y, Marchal, Paul. A spatial model of the mixed demersal fisheries in the Eastern Channel. In: Marine Productivity: Perturbations and Resilience of Socio-ecosystems Proc 15th French-Japan Oceanogr Symposium. H.-J. Ceccaldi *et al.*; 2015. p. 187-95.

Ecoregion H: Western Mediterranean Sea

There is no progress to report on multispecies modelling in the Ecoregion this year.

Ecoregion I: Adriatic-Ionian Seas

There is no progress to report on multispecies modelling in the Ecoregion this year.

Ecoregion J: Aegean-Levantine

There is no progress to report on multispecies modelling in the Ecoregion this year.

Ecoregion K: Oceanic northeast Atlantic

There is no progress to report on multispecies modelling in the Ecoregion this year.

Ecoregion L: Baltic Sea

Food dependent growth

Five decades of stomach content data allowed detailed insight into the long-term development in consumption and diet composition, and the resulting somatic growth of Atlantic cod in the Baltic Sea. Small cod feed almost exclusively on benthos and food availability in this benthic phase has been driven by the environment. We show a recent reversal in the ontogenetic development of feeding level over body length. Present feeding levels of pre-spawning cod imply severe growth limitation and increased starvation-related mortality (Fig. 1), which extends the post-larval settlement bottleneck and blocks the transition to piscivory. The low growth rate and high mortality rate of the young cod manifest as a reduction in size-at-age (Fig. 2) and low abundance which are determined early in life. These results suggest that density-dependence is environmentally mediated and hence not stable under environmental change (Neuenfeldt *et al.* in prep.).

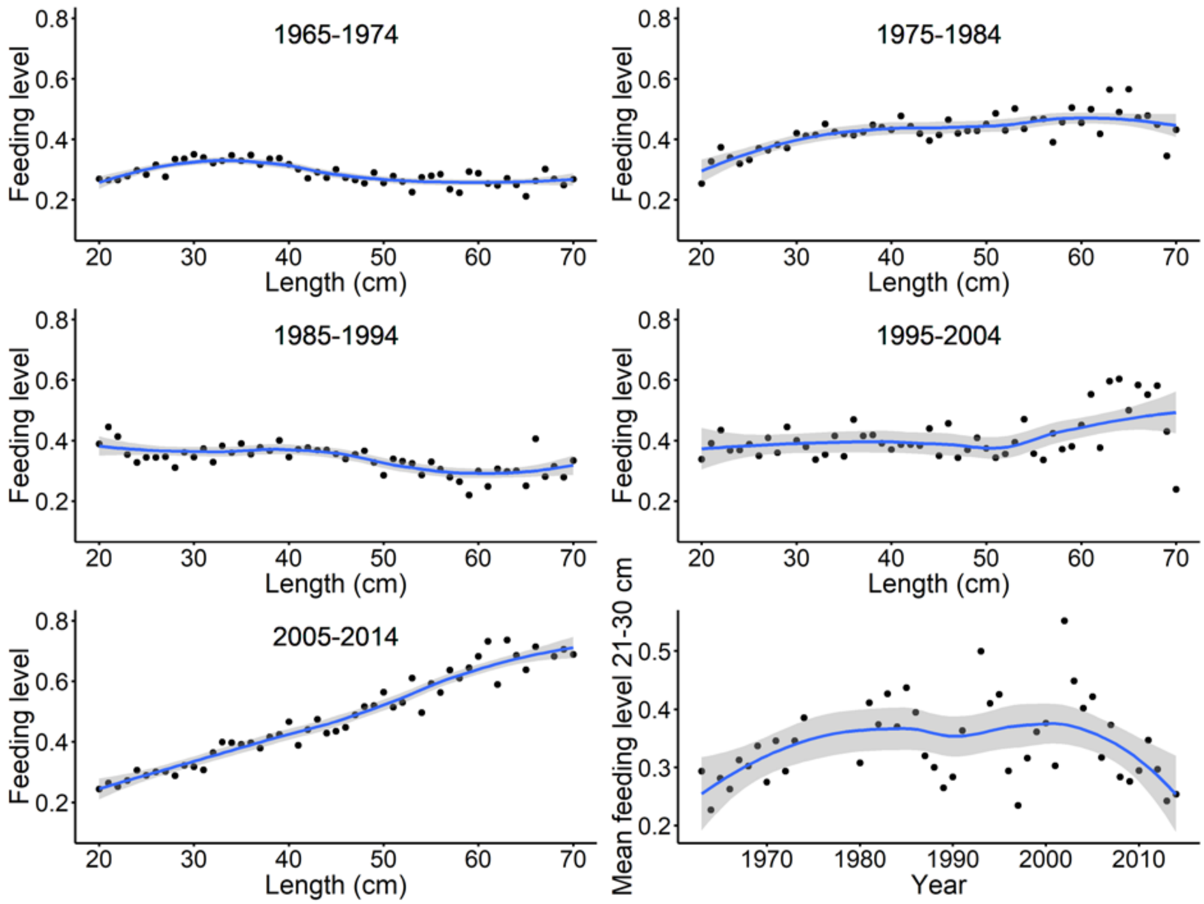


Figure 1. Feeding levels over *G. morhua* length during the past five decades. LOESS-based smoothed trends are plotted in blue together with shadowed confidence limits.with The lower right panel: feeding level over time for *G. morhua* of 21–30 cm total length.

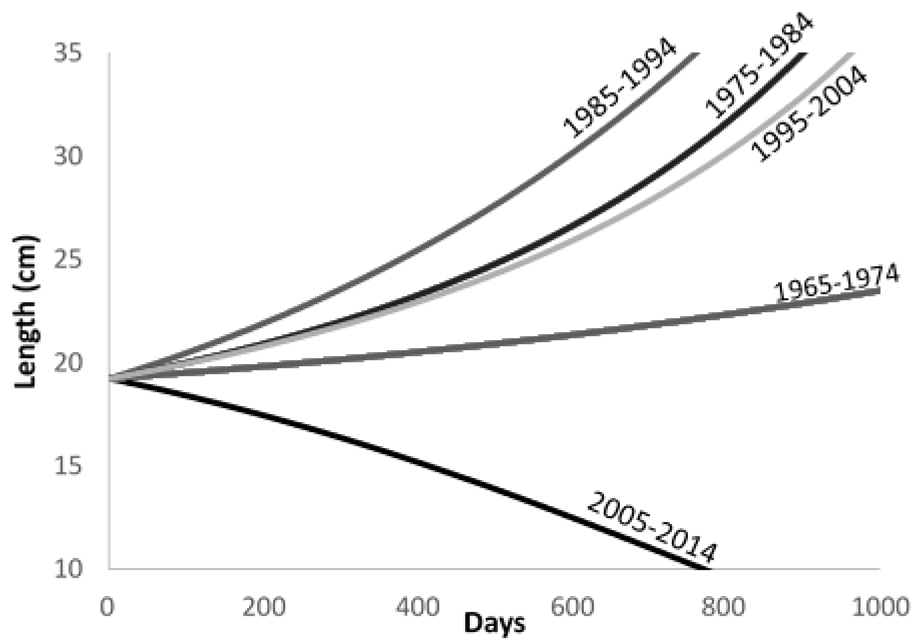


Figure 2. Simulated *G. morhua* growth trajectories for the five decades covered by the stomach sampling programme. Each simulation was started at 20 cm total length and stopped at 35 cm.

Cod cannibalism

A large database of cod stomach content data is available for the Baltic Sea, covering the period from 1960s–2010s. The stomach data have been collected and worked up within a number of national and international projects over the past decades, some of the most recent efforts with this database were carried out within EU BONUS INSPIRE project.

This database was used to address a specific question related to cod cannibalism as a source of natural morality. Given the poor nutritional condition of EB cod in later years, which indicates food limitation, it is relevant to consider whether cannibalism related mortality has subsequently increased. Historical estimates of predation mortalities related to cod cannibalism back to 2011 are available from multispecies model (SMS). The estimates from this model have however not been possible to update for more recent years due to lack of age-based assessment for EB cod, which is needed for the former SMS model. Analyses of stomach data alone show that frequency of occurrence of cod in stomachs was high in years in the period 2007–2012 (Fig. 1).

The results suggest that cannibalism related mortality was relatively high around 2010, however has substantially declined in later years (Fig. 2). This is likely related to low abundance of larger individuals in the cod stock that can use cod as prey. It should be noted that the exact values of natural mortality from the recent estimates are relatively more uncertain than the relative dynamics, which we mainly focus on. These results suggest that cannibalism could have contributed significantly to natural mortality of smaller cod in early 2010s, however presently the cannibalism related mortality seems relatively low (Neuenfeldt *et al.* in prep.)

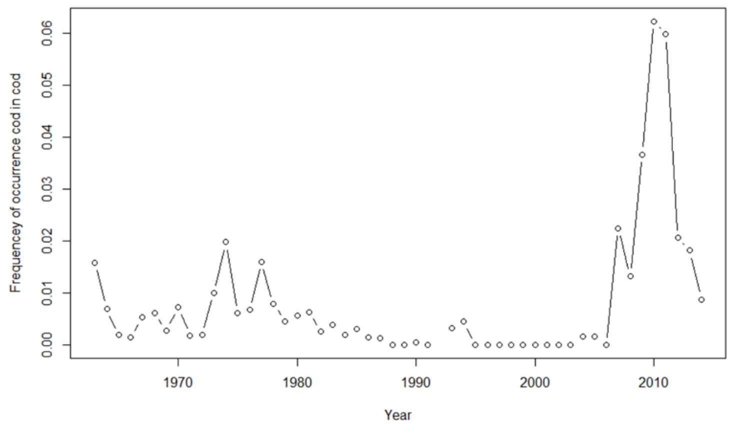


Figure 3. Frequency of occurrence of cod in cod stomachs (Neuenfeldt *et al.* in prep).

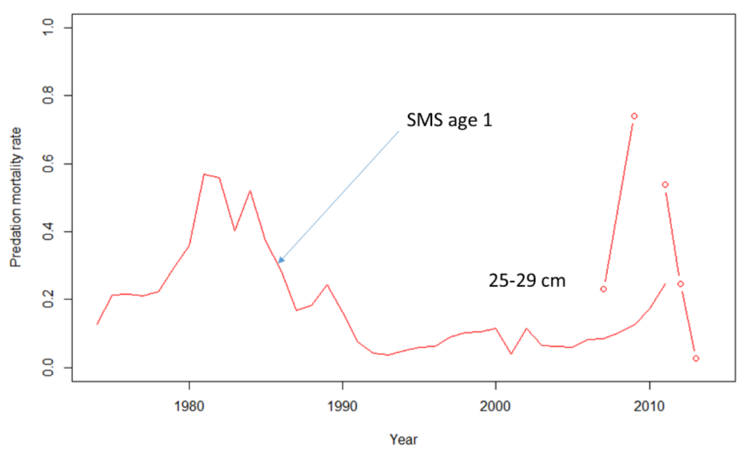


Figure 4. Estimates of predation mortality for age 1 cod from SMS model until for 1974–2011, compared to the recent estimates for 25–29 cm cod, based on BITS survey and new stomach data (Neuenfeldt *et al.* in prep).

Multi-species food web model for the Kattegat-Skagerrak area: a new perspective towards an Ecosystem-Based Fishery Management (EBFM)

Maciej T. Tomczak, Andrea Belgrano

There is currently a knowledge gap in ecosystem functioning and species interactions using food webs and an ecological networks perspectives for Kattegat and Skagerrak. The work in progress is to complete an existing Ecopath-EwE multi-species trophodynamics model for the Kattegat, and to extend the model to Skagerrak.

Kattegat Ecopath EwE Model description ver. 0.1

The mass-balance trophic model describes the annual conditions in the Kattegat ecosystem in 1984 based on a large amount of available historical information. The

Kattegat food-web model includes 31 functional groups (Fig. 1) with trophic components from primary producers to top predators. The human impact was indicated by using fishing effort for metier related fishing fleets as: gillnetter, set of gillnets, tarammel nets, bottom otter trawls, pelagic otter trawls, bottom pair trawl, anchored seine. The Ecopath food-web model for 1984 will be updated and extended to Skagerrak. The temporal part of the modelling - Ecosim is calibrated for the period 1984–2005 using observed time-series from almost all trophic levels, and include various type of extensive drivers as fishing, eutrophication and climate change. The model perform relatively well and it's able to partly reproduce observed dynamic.

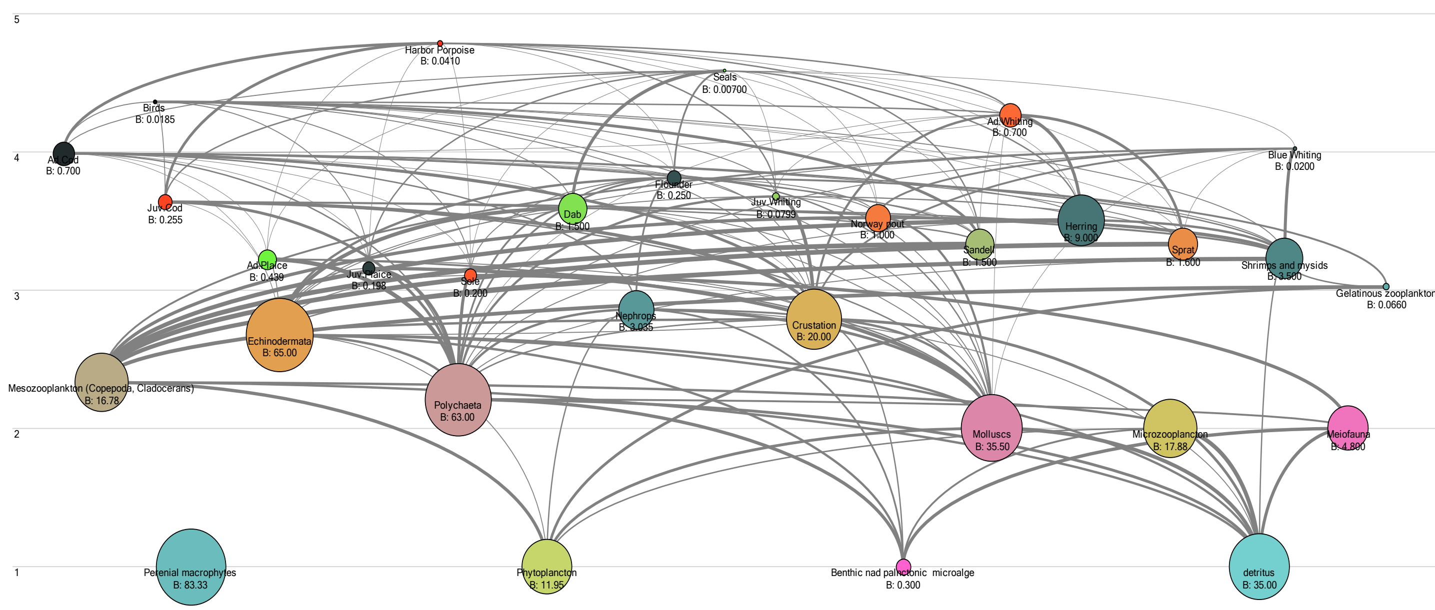


Figure 2. Kattegat Ecopath EwE model food-web structure 1984.

Ecoregion M: Black Sea

There is no progress to report on multispecies modelling in the Ecoregion this year.

Ecoregion: Canadian Northwest Atlantic

The Northwest Atlantic Fisheries Organization (NAFO), on its amended convention, states the aim of applying the ecosystem approach to fisheries management, taking into account the interactions between the different components in the ecosystem. The European Union, through the SC05 project *"Multispecies Fisheries Assessment for NAFO"*, contributes to the development of the multispecies approach in the NAFO area, with the Flemish Cap cod, redfish and shrimp fishing system (NAFO area 3M) as a case study. As part of the tasks of this project, as a first goal the multispecies model GadCap (Pérez-

Rodríguez *et al.* 2017) has been updated, extending the model time coverage up to 2016. Some components of the model have been improved, like those defining biological processes like growth or maturation, trophic interactions and fishing fleets. The effect of fishing, trophic interactions (including cannibalism) and water temperature in the dynamic of these three major fishing resources has been modelled. The model highlights the interdependent dynamic of these stocks, and reveals strong interactions between recruitment, fishing and predation (including cannibalism). These drivers have shown marked changes in their relative importance by species, age, and length over time, producing a transition from a traditional redfish-cod dominated system in the early 1990s, to an intermediate shrimp-other fish species state by late 1990s, and in turn back to something close to the initial state by late 2000s. Hence, the results of the model clearly indicate that the dynamic of the Flemish Cap cod, redfish and shrimp stocks is strongly interconnected, with predation mortality being a main driver of changes in population structure and biomass over time, and very relevant the role of cannibalism in cod and redfish.

Ecoregion: US Northwest Atlantic (and other regions)

There are updates to several models ongoing in the Northeast US; other work is reported under ToRs C, D, and G.

Implementing MSE using Rpath as an operating model (Contributed by Sean Lucey, Sean.Lucey@noaa.gov)

Until recently, full feedback interactions between a management strategy and an EwE-based operating model were impractical. However, with the development of Rpath, an R implementation of the EwE algorithms, users now have the ability to fully customize the operating model to be conditioned on the outputs from external assessment models. Rpath now has the capability to pause after every time step, evaluate an external model, and receive information back from that external model which can modify parameters that effect the next time step. This gives the user an opportunity to evaluate a range of management strategies in an ecosystem context. This ability was demonstrated using an existing Georges Bank model. Three harvest strategies were tested by simulating 100 years of data. For each time step, survey data were generated from Rpath by adding lognormal error to biomass outputs. This data was then input to a simple surplus production model (discrete Schaefer model) to replicate an assessment every fifth year. The results of which are passed to harvest control rules that mimic characteristics of management for commercially targeted species on Georges Bank. Benefits of particular strategies on the target species are evaluated, along with impacts to other parts of the ecosystem. Therefore, the evaluation criteria for selecting a management strategy can be based on the inherent trade-offs within the system. For this example, all three strategies produced on average the MSY for the target species but two strategies were better for a secondary “choke” species (Fig 1). If this had been an actual MSE process, the managers would more than likely chose the strategy that did not negatively effect the “choke” species while at the same time trying to minimize the variability in catch. These types of decisions are not possible when dealing within a single species context.

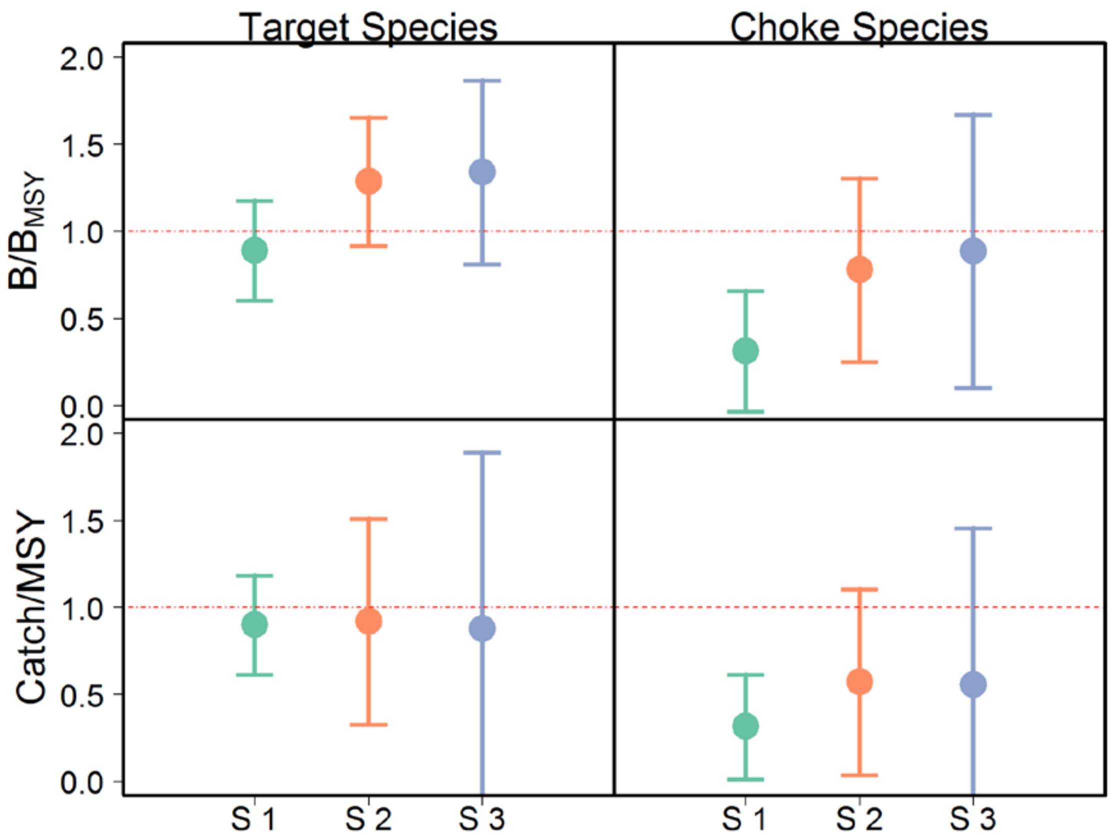


Figure 1. Rpath MSE demonstration results.

Menhaden ecological reference points and multispecies modelling (contributed by Howard Townsend, Howard.Townsend@noaa.gov):

The US Atlantic States Marine Fisheries Commission (ASMFC) Atlantic Menhaden Management Board (Board) established a working group (WG) tasked with developing ecological reference points (ERPs) for menhaden using multispecies/ecosystem models.

The WG has developed a suite of novel multispecies models to ensure they are able to generate ERPs which meet as many management objectives as possible. The models being developed include:

a) Bayesian surplus production model with time-varying population growth rate – This models estimates the trend in total Atlantic menhaden stock biomass and fishery exploitation rate by allowing the population growth rate to fluctuate annually in response to changing environmental conditions. This approach produces dynamic, maximum sustainable yield-based ecological reference points that implicitly account for the forage services menhaden provide.

b) Steele-Henderson model (i.e., surplus production with forced biomass of predators) – This model permits non-fisheries (predation and environmental) effects to be quantified and incorporated into the single species stock assessments, allowing fixed and non-equilibrium (time-varying) ecological overfishing thresholds to be established. This approach is not intended to replace more complex multispecies ecosystem assessment models, but rather to expand the scope of the single species assessments to include the separate and joint effects of fishing, predation and environmental effects at the fish community level.

c) Multispecies statistical catch-at-age model (MS-SCAA). This model uses standard statistical catch-at-age techniques and single species models are linked using trophic calculations to provide a predator-prey feedback between the population models.

d) Ecopath with Ecosystem models. The model is flexible and able to explore additional menhaden relevant scenarios, ERPs, and questions. Two version of the model are being developed. The WG will use a highly articulated EwE model (61 tropic groups and 8 fleets) adapted from Buchheister *et al.* (2017 a & b). This model will be used to evaluate the other models and erps being developed. The WG will also use simpler model focused on menhaden and the few predators modelled in the MS-SCAA.

These models are being updated and reviewed during 2019. Once these models are fully vetted by the WG, the WG will select a set of models for peer-review in November 2019 along with the single-species stock assessment model, which has traditionally been used for menhaden management.

Northwest Atlantic Continental Shelf (NWACS) EwE model (contributed by Andre Buchheister, andre.buchheister@humboldt.edu)

Objectives

- Evaluate the ecosystem impacts of different Atlantic Menhaden fishing mortality rates to help inform fisheries management of Menhaden
- Compare the performance (in an ecosystem context) of potential reference points proposed for Atlantic menhaden management

EwE Model

- Single spatial domain (NE USA)
- Years: 1982–2013; project forward 50 yrs
- 8 fishing fleets; 61 trophic groups (Fig 2)
- Performance metrics: Biomasses, Yields, Ecosys. Structure

Objectives

1) Update the NWACS model to enhance utility for current management (add data through 2017, add primary production forcing, improve model structure, etc.)

2) Evaluate ecosystem tradeoffs and predator impacts that result from using single species Menhaden reference points & alternative Ecological Reference points (ERPs)

3) Use management strategy evaluation to examine management and reference point effectiveness in the presence of uncertainties

Status

Project is underway; funded by Lenfest Ocean Program

NWACS model being considered for use as strategic tool for reference point assessment

Timeline:

Objectives 1 and 2 – summer/fall 2019 (in conjunction with Menhaden Assessment)

Objective 3 – Aug 2020

Contact

Andre Buchheister, Department of Fisheries Biology, Humboldt State University andre.buchheister@humboldt.edu , 707–826–3447

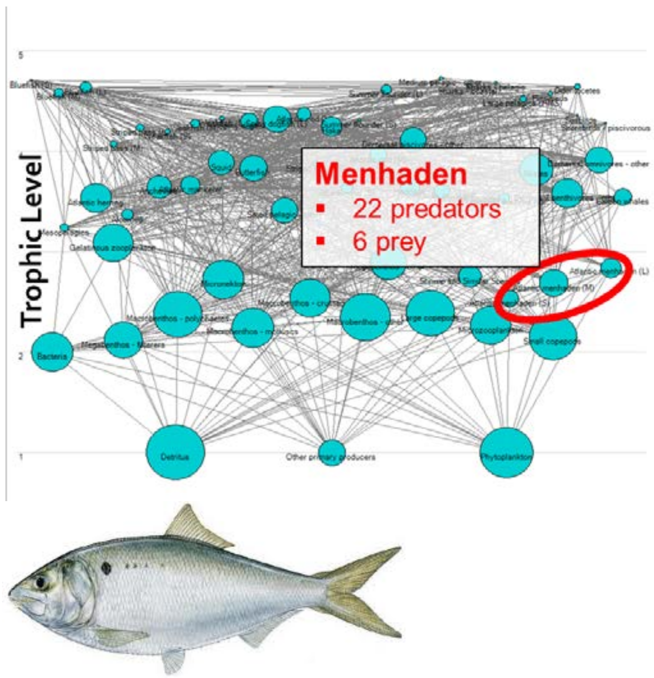


Figure 2. EwE model used for Atlantic menhaden simulations.

References

Buchheister, A., Miller, T. J., Houde, E. D. (2017a). Evaluating ecosystem-based reference points for Atlantic menhaden (Brevoortia tyrannus). Marine and Coastal Fisheries. doi: 10.1080/19425120.2017.1360420.

Buchheister, A., Miller, T. J., Houde, E. D., and Loewensteiner, D. A. (2017b). Technical Documentation of the Northwest Atlantic Continental Shelf (NWACS) Ecosystem Model. Report to the Lenfest Ocean Program, Washington, D.C. University of Maryland Center for Environmental Sciences Report TS-694–17. Available at: http://hjort.cbl.umces.edu/NWACS/TS_694_17_NWACS_Model_Documentation.pdf .

US West Coast Atlantis modelling applications (contributed by Isaac Kaplan, Isaac.Kaplan@noaa.gov)

For the California Current ecosystem on the US West Coast, Marshall *et al.* (2017) projected impacts of ocean acidification on the full food web. This involved downscaled oceanographic projections under IPCC climate scenario RCP8.5, including pH, via a Regional Ocean Modelling System (ROMS). The biological response to ocean acidification was informed by a meta-analysis of over 300 experimental studies reporting sensitivity of organisms to pH (Busch and McElhany 2016). The ROMS oceanography and the understanding of direct pH impacts were then used to evaluate indirect, food web effects using an Atlantis ecosystem model. Results suggest that greater exposure to more acidic water leads to stronger declines of most invertebrates, such as sea urchins, in the southern regions (Hodgson *et al.* 2018; Marshall *et al.* 2017). Loss of prey items drove declines in biomass and revenue of some groundfish species, and also contributed to declines in Dungeness crab. Hodgson *et al.* (2018) translated this loss of revenue due to OA into economic impacts and employment at the port level. This involved using an economic input-output model to translate from dockside revenue to jobs and income impact in the broader economy. Results illustrate that Dungeness crab is a major player: northern ports heavily reliant on Dungeness crab are projected to experience the largest losses of revenue, income, and employment, even though some other species declined more strongly in the south (but were less economically important; Fig 3).

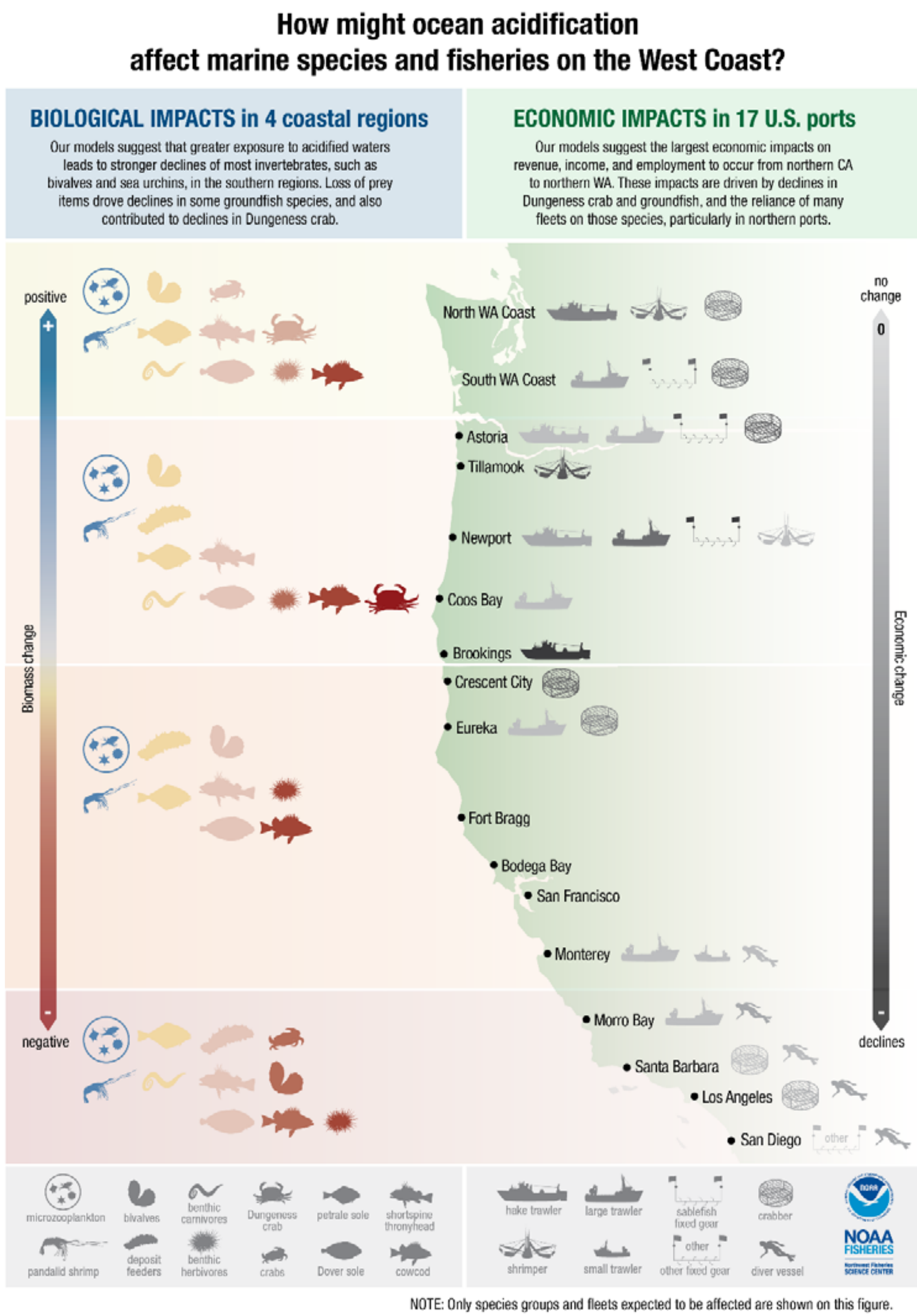


Figure 3. Results from OA simulations, impacts to fishing ports on the US West Coast.

References

Busch, D. Shallin, and Paul McElhany. 2016. "Estimates of the Direct Effect of Seawater pH on the Survival Rate of Species Groups in the California Current Ecosystem." PloS One 11 (8): e0160669.

Hodgson, Emma E., Isaac C. Kaplan, Kristin N. Marshall, Jerry Leonard, Timothy E. Essington, Shallin D. Busch, Elizabeth A. Fulton, Chris J. Harvey, Albert Hermann, and Paul McElhany. 2018. "Consequences of Spatially Variable Ocean Acidification in the California Current: Lower pH Drives Strongest Declines in Benthic Species in Southern Regions While Greatest Economic Impacts Occur in Northern Regions." Ecological Modelling 383 (10): 106–17.

Marshall, Kristin N., Isaac C. Kaplan, Emma E. Hodgson, Albert Hermann, D. Shallin Busch, Paul McElhany, Timothy E. Essington, Chris J. Harvey, and Elizabeth A. Fulton. 2017. "Risks of Ocean Acidification in the California Current Food Web and Fisheries: Ecosystem Model Projections." Global Change Biology 23 (4): 1525–39.

Northeast US Atlantis update (contributed by Ryan Morse, Ryan.Morse@noaa.gov)

Work completed:

- New group structure 62-> 89 groups
- New hydrodynamic forcing: exchanges, S, and T from ROMS DOPPIO model
- Updated initial conditions N values based on new length at age parameters
- Updated distributions, growth, grazing dynamics for all groups
- New diet matrix tuned to consumption

Work in progress:

- Fleet structure, fishing module in process of being updated
- Force lower trophic level groups with time-series data to mimic regime shifts, test effect on fish production
- Test sensitivity to hydrodynamics - effects of hydrodynamic model selection (data-assimilative Roms vs non-DA, ROMS DOPPIO vs HYCOM)

Northeast US multispecies model comparisons (contributed by Jason Boucher, Jason.Boucher@noaa.gov)

Objective

To assess the impact of structural uncertainty in length- and age-based multispecies population dynamic models on the estimation of underlying population parameters using biomass, recruitment, mortality, and predation as metrics

Models

Multispecies Statistical **Catch-at-Age** Model (Curti *et al.* 2013)

Hydra Multispecies Statistical **Catch-at-Length** Model (Gaichas *et al.* 2017)

Status

Catch-at-Age Model is operational (Figure 4)

Length-at-Age Model is almost operational

Completing final likelihood component - predation

Awaiting simulated data from Atlantis model to run 'known' conditions through the multispecies models

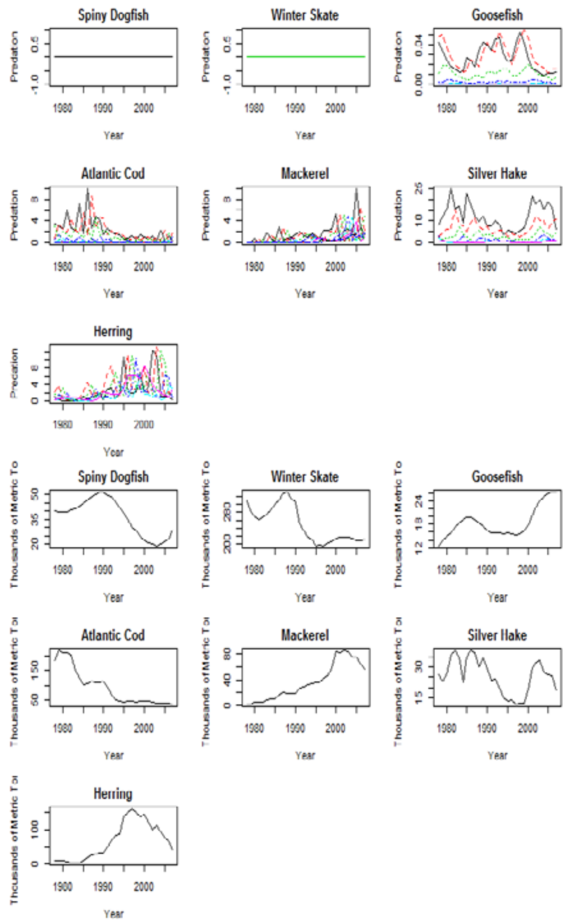


Figure 4. Preliminary catch at age model total biomass and predation at age estimates.

References

Curti, K. L., Collie, J. S., Legault, C. M., Link, J. S., and Hilborn, R. 2013. Evaluating the performance of a multispecies statistical catch-at-age model. Canadian Journal of Fisheries and Aquatic Sciences, 70: 470–484.

Gaichas, S. K., Fogarty, M., Fay, G., Gamble, R., Lucey, S., and Smith, L. 2017. Combining stock, multispecies, and ecosystem level fishery objectives within an operational management procedure: simulations to start the conversation. ICES Journal of Marine Science, 74: 552–565.

Applying CEATTLE to GOA groundfish (contributed by Grant Adams, adamsgd@uw.edu)

- ADMB based MSCAA model developed for groundfish in the Bering Sea (BSAI), contributed last year by Holsman.
- 3 age-structured models linked by bioenergetics based predation mortality

    - MSVPA based suitability
- Parameterized to incorporate temperature forcing
- Supplement to current SAF

Current efforts:

- Build CEATTLE in TMB
    - Currently multispecies version implemented in TMB for BSAI
    - Modularize and make flexible
        - Diet estimation
    - Benefits:
        - Random effects, faster estimation, flexible, relative ease
- Apply CEATTLE to GOA groundfish (Fig 5)
- Synthesize and incorporate GOA data
    - Estimate parameters for pacific cod, Pollock, arrowtooth flounder, and halibut

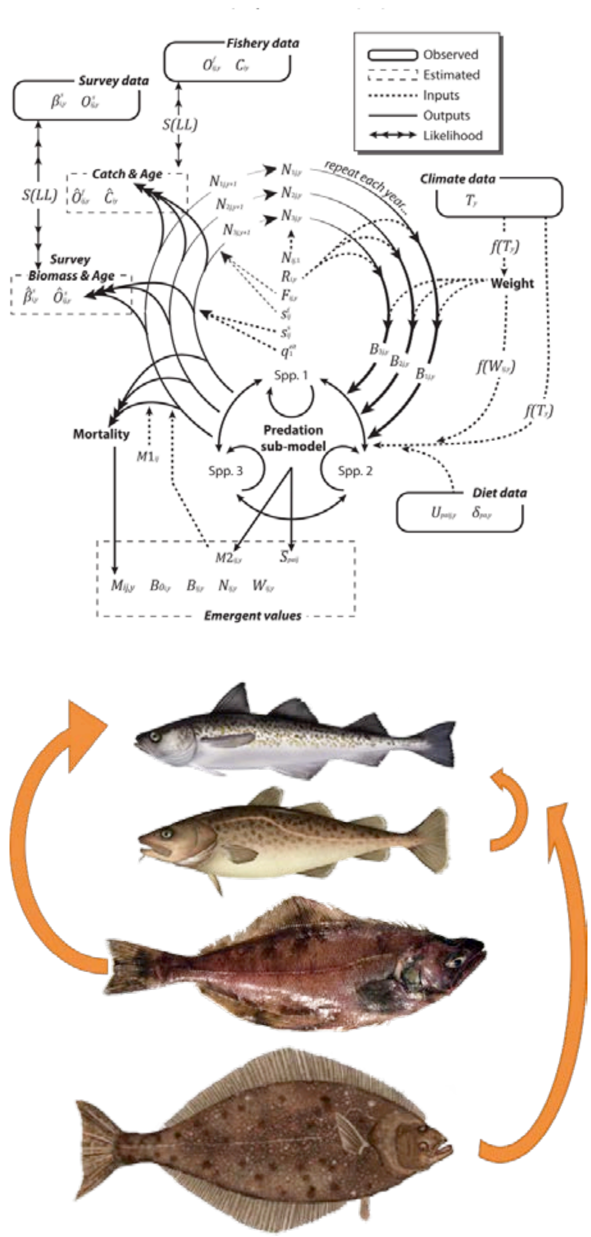


Figure 5. Holsman *et al.* CEATTLE model framework, and species included for the GOA.

References

Holsman, K. K., Ianelli, J., Aydin, K., Punt, A. E., and Moffitt, E. A. 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.

Other US multispecies and ecosystem updates

Updates to the Herring MSE for New England Fishery Management Council, fully described in 2017 WGSAM report

- Paper available: Deroba, J. J., Gaichas, S. K., Lee, M.-Y., Feeney, R. G., Boelke, D. V., and Irwin, B. J. 2018. The dream and the reality: meeting decision-making time frames while incorporating ecosystem and economic models into management strategy evaluation. Canadian Journal of Fisheries and Aquatic Sciences. http://www.nrcresearchpress.com/doi/10.1139/cjfas-2018–0128.
- The Council picked a control rule! September 25, 2018, see press release.

Updates to the Mid-Atlantic EAFM process, fully described in 2016 and 2017 reports

- Indicator based risk assessment completed (paper in review Frontiers Marine Sci.)
- Next step: develop conceptual model integrating climate, habitat, food web, fishery, economic, and social effects
- Council selected summer flounder as a focal species
- Could lead to MSE one year out, will need all your models, stay tuned

Annex 6: ToR C. Consider methods to assess the skill of multispecies models intended for operational advice

Performance of multispecies assessment models: insights on the influence of diet data

For summary:

A simulation analysis was performed to investigate how diet data quality and availability and the method of fitting to diet data could influence multispecies assessment models (Trijoulet *et al.* in review). Four operating models that simulate trophic interactions for two fish species and different scenarios of diet data availability or quality. The simulated data sets were fitted using 4 statistical catch-at-age models that estimated fishing, predation and residual natural mortality and differed in the way the diet data was fitted.

- Fitting the models to diet data averaged overtime should be avoided since it resulted in estimation bias.
- Fitting annual diet composition per stomach presented bias estimates due to the occurrence of zeros in the observed proportions and the statistical assumptions for the diet model.
- Fitting to annual stomach proportions averaged across stomachs led to unbiased results even if the number of stomachs was small, the interactions were weak or some sampled years were missing. These methods should be preferred when fitting multispecies models.

Main text:

The simulation study presented here comprised 4 operating models (OMs) and 4 estimation models (EMs). The multispecies models were developed using the R package Template Model Builder (TMB). The OMs are identical in all aspects except for how predator diet data are simulated.

- OM1 represented an "ideal" situation where the number of samples was large (500 stomachs per year) and the predator-prey interactions were strong such that the modelled prey represented on average 40% of the predator diet and around 50% for older predator age groups. OM1 was used as a base case and other OMs were created by varying from this base case.
- The importance of the number of diet samples was evaluated by using only 50 (the first 10%) of the stomachs in OM2.
- The effect of the strength of the predator-prey interactions on estimation was assessed with OM3 where predator-prey interactions were reduced by increasing the biomass of other food available such that it represented around 80% of the predator stomach contents for all simulated data sets.
- Often, diet data are not available for the entire time-series. For simplicity, in OM4, we assumed that diet data were not available for the first half of the mod-

elled time-series, but predation was still modelled over the entire model time frame.

Given assumed parameter and input values, 1000 sets of stochastic observations were generated for each OM. The simulated data sets included aggregated catch and abundance indices, age composition data for catch and abundance indices, and predator diet data.

Four estimation models (EMs) were fitted to the 4 OMs. The EMs are all multispecies statistical catch-at-age models that differ in how the diet data inform the model.

- EM1 represented the same configuration as the OMs. Each diet observation (stomach) was fitted.
- EM2 was a simplification where diet observations were averaged over stomachs for predators of a specific age, resulting in only one data point per predator age and year.
- In EM3, the model was simplified further by averaging diet observations and corresponding predictions for each predator age over 10 year intervals. Aggregating diet over time is sometimes used when the number of annual samples is deemed insufficient.
- Finally, EM4 predicted and fitted to the mean diet proportions over the entire time-series. Performance of models EM3 and EM4 should provide insight on the effect of aggregating the diet data over time.

Main conclusions:

- Fitting the models to diet data averaged overtime should be avoided since it resulted in estimation bias.
- Fitting annual diet composition per stomach presented bias estimates due to the occurrence of zeros in the observed proportions and the statistical assumptions for the diet model.
- Fitting to annual stomach proportions averaged across stomachs led to unbiased results even if the number of stomachs was small, the interactions were weak or some sampled years were missing (Figure 1). These methods should be preferred when fitting multispecies models.

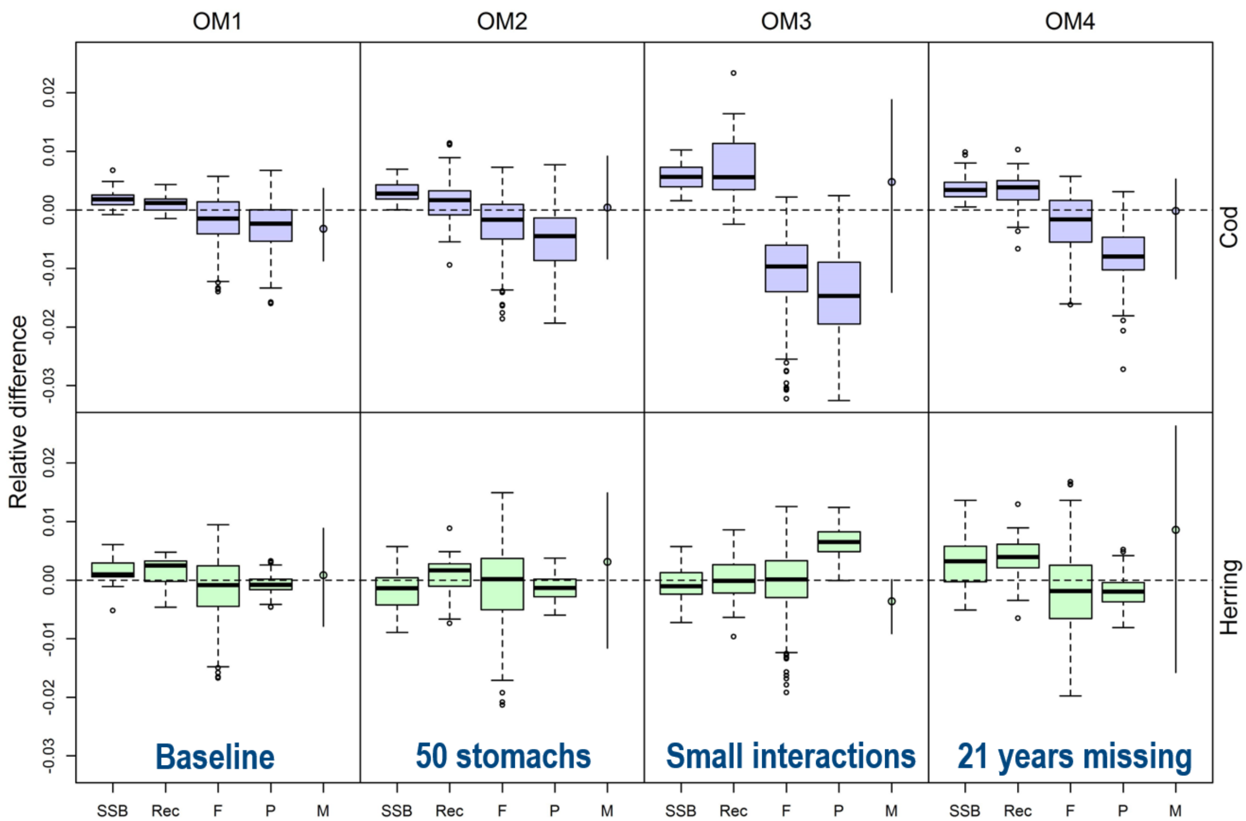


Figure 3. Median relative differences for spawning stock biomass (SSB), recruitment (rec), fishing mortality (F), predation mortality (P) and residual mortality (M) for the EM2 fitted to all OMs. One value is estimated for residual natural mortality (M), so the median relative difference and its 95% confidence interval are presented for M.

For more details: Vanessa Trijoulet vtri@aqua.dtu.dk

Trijoulet, V., Fay, G., Curti, K., Smith, B., and Miller, T. J. in review. Performance of multispecies assessment models: insights on the influence of diet data. ICES Journal of Marine Science.

Skill Assessment of multispecies/ ecosystem models and „Keyruns" – Examples from WGSAM (Alexander Kempf *et al.*)

- **Various examples of skill assessments were carried out by WGSAM experts in recent years. Different people focus on different aspects. There is a need to come up with best practice guidelines**
- **Testing just the hindcast performance is not sufficient for management. Predictive skills have to be tested. Model ensembles are an important way to improve prediction skills.**
- **Keyruns can be conducted during WGSAM and it may be the best place to evaluate multispecies and ecosystem models as the expertise on such models is often limited at "single species" benchmark meetings.**

An overview on the work done in WGSAM during the last years on skill assessments and Keyruns has been presented. This is a basis to summarise the work under TOR c. There is an urgent need for proper skill assessments and benchmarks to strengthen the trust in the output from complex models. By doing this, skill assessments have to be carried out for hindcasts and forecasts separately and on the right scales dependent on the questions asked for advice. Different examples from literature with members of WGSAM involved highlighted different ways to conduct skill assessments with different skill metrics and focus on different aspects (e.g., sensitivity analysis, metrics to compare hindcasts and forecasts to observational data, Prebal for EwE models).

Examples from the work of WGSAM members in the last two years included analysis on the prediction skills of diet selection models (from Nataliia Kulatska) as well as a study in Icelandic waters to evaluate the performance of EwE using Atlantis as operating model (from Erla Sturludottir and Gunnar Stefansson). The latter study highlighted the fact that the performance of hindcasts does not allow to judge on predictive skills. A study from Gaichas *et al.* tested the performance of three different multi species models and multi-model inference in an MSE type approach. The conclusion was that the model ensemble outperformed individual models with realistic input data conditions (i.e. uncertainty and bias in input data).

Keyruns are a core activity of WGSAM and they refer to a model parameterization and output that is accepted as a standard by ICES WGSAM, and thus serves as a quality assured source for scientific input to ICES advice. The importance of detailed documentation of input, model settings and diagnostics has been highlighted. Output has to be presented in an easy accessible format for other working groups and people to allow an efficient use of model results. WGSAM uses Github, an extra stock annex on the ICES website, standardized main output (tables and figures) and puts effort into the direct communication with e.g., assessment working groups.

Conclusions:

There are various ways of testing the skills of models in the literature and various examples can be found in ICES groups. However, different people focus on different aspects. Therefore, there is a need to come up with best practice guidelines for different types of models to establish standards in ICES before a model can be used for advice. Skill assessments for hindcasts only may not be sufficient (but depends on the questions asked). For decision making prediction skills are often important but are less frequently tested. This needs to be changed. Model ensembles are an interesting way to improve the prediction skills compared to using one particular model only. Keyruns can be conducted during expert group meetings. However, dedicated members are needed who work intersessionally. Keyruns are not there to replace a peer review of the modelling method itself. It needs to be discussed whether extra benchmark meetings are needed or whether "Keyruns" during expert working groups are a better alternative given a potentially higher participation of experts working with complex models.

Global sensitivity analysis of a multi-species size spectrum model of the North Sea

A derivative-based global sensitivity analysis was applied to the North Sea multispecies version of the *mizer* size spectrum model. 307 parameters and 7 model outputs, including population size, biomass, spawning stock biomass, the large fish indicator, mean weight, and the community slope, were considered in the analysis. Each parameter was assumed to follow a uniform distribution with an upper and lower limit of ±10% of the nominal parameter value. The sensitivity indices were presented for community biomass as an example of the results. The community biomass was most sensitive to the parameters associated with fishing, the size of the background resource (i.e. other food sources such as zooplankton and phytoplankton), and the metabolic, search, and feeding rates of the 12 fish species included in the model. The community biomass was least sensitive to the initialisation parameters (i.e. the initial population size of each of the 12 fish species and the starting slope of the community spectrum), and the parameters associated with recruitment and background mortality. Further research is needed to apply more realistic distributions to each of the parameters to identify areas in which to focus data collection to reduce the uncertainties in the model outputs.

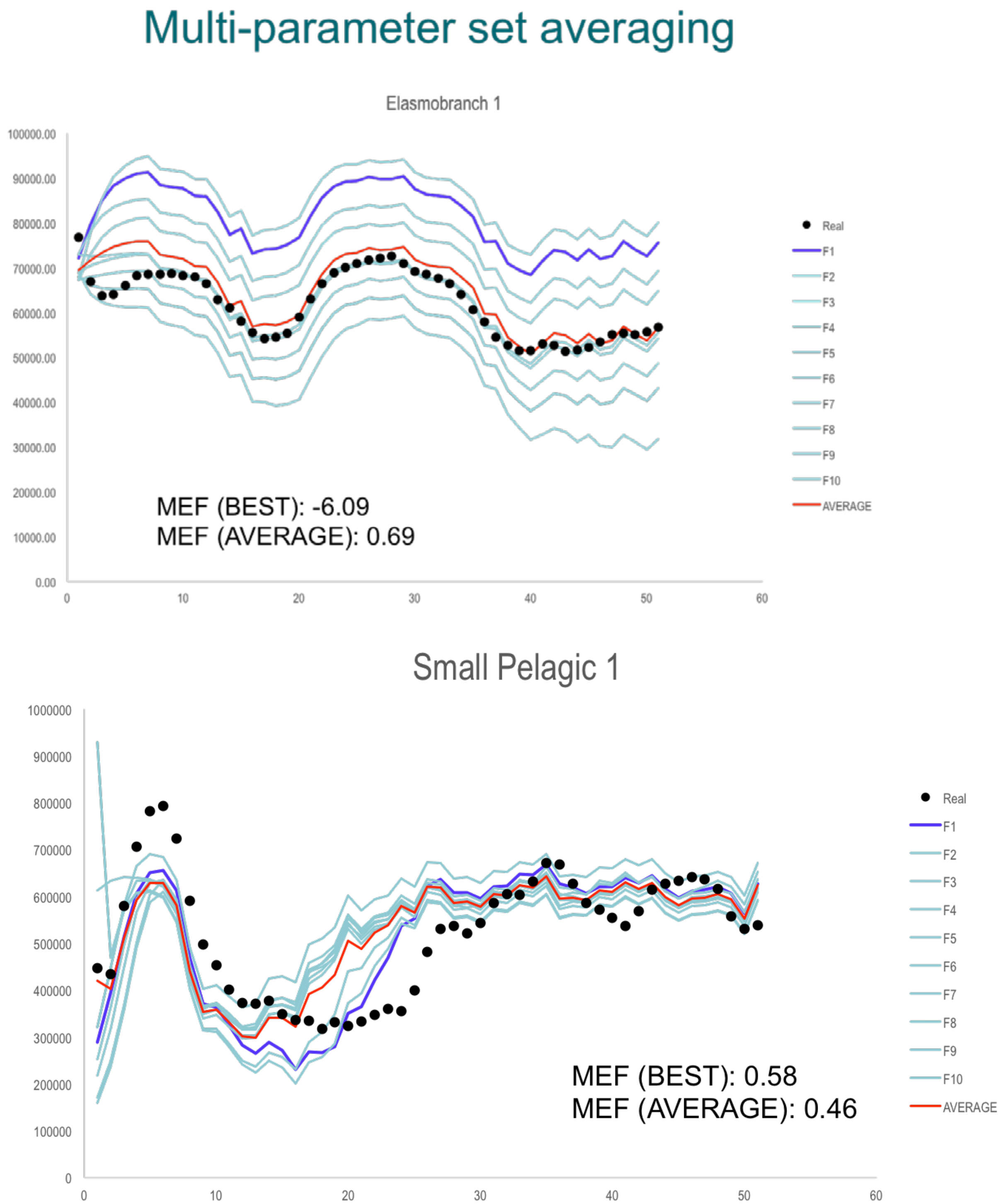
Annex 7: ToR D. Investigate the performance of multi-model ensemble in comparison to single model approach

One Model – Many Parameters: A Multi-Parameter Inference Framework (contributed by Robert Gamble, Robert.Gamble@noaa.gov)

The Kraken Multispecies Production Modelling software package was used to take first steps towards exploring a multi-parameter inference framework in much the same fashion as multi-model inference can be done. The basic concept is that if the fitness landscape has multiple local minima in addition to a global minima, starting from different sets of parameter values will result in different estimate parameters given a directed search using traditional estimation methods. Additionally, if the parameters chosen are within a local minima rather than the global minima, there is no chance under most estimation methods currently used to find the global minima instead. Informed expert knowledge can suggest starting parameters, but in many cases those parameters are not based on empirical or laboratory studies and could have a large range of plausible values (competition, other mortality, predation, etc.).

The operating model for our analysis was Hydra, a length structured multispecies model with explicit recruitment and predation. The specific operating model simulated 10 species of fish. By contrast, the multispecies production model included explicit predation but only the usual growth rate for the entire population. Random white noise error was applied to both the biomass and the catch outputs from Hydra, and a genetic algorithm in Kraken was used to fit the parameters. A genetic algorithm searches the parameter space by starting with random parameter sets within specified bounds, calculating the fitness of each set of parameters, and creating offspring models from the previous generation of models through recombination and mutation. While fitter models are more likely to be selected to create the offspring, any previous model could be selected as a parent.

The top 10 models based on the Modelling Efficiency Factor (MEF); (Stow *et al.* 2009) objective function were then plotted against the true biomass curves from Hydra, and the MEF calculated for the whole model and each species for each of the 10 best parameter sets. Additionally the average biomass of the 10 best parameter sets was calculated for each species and also had the MEF calculated for each species. Our analysis focused mostly on the top predator (Elasmobranch 1) and top prey (Small Pelagic 1) in the system. While Elasmobranch 1 had a better MEF using the best parameter set than taking the average of all 10 best models, it still had a good MEF (greater than 0) using the average. By contrast, Small Pelagic 1 had a very poor MEF (lower than 0) in the best overall parameter set, but had a very good MEF using the average. This method will be explored further in order to determine whether the initial results hold true across more species and under different types of parameterizations.



Stow, C.A., Joliff, J., McGillicuddy, D.J., Doney, S.C. Allen, J.A., Friedrichs, M.A.M, Rose, K.A., Wallhead, P. 2009. Skill Assessment for Coupled Biological/Physical Models of Marine Systems. J. Mar. Syst. 76(1–2): 4–15.

Testing multi-model methods for the US Atlantic Herring stock assessment (contributed by Sarah Gaichas Sarah.Gaichas@noaa.gov)

Here, we apply a simple MMI approach to two herring stock assessment models. The objective is to "field test" a simple method during a working group meeting so that we can see what would be necessary to evaluate an MMI approach during a benchmark assessment process. We then make suggestions to build a framework for the use of MMI in the future for providing advice to fishery managers.

Methods

For the purposes of this example, there was no opportunity to develop a set of candidate models based on alternative structural hypotheses (Burnham and Anderson 2002), although this happens iteratively through the working group process within a single model framework as alternative model configurations and dataset combinations are discussed, tested, and accepted or rejected as improvements on a baseline model. We consider a "model" to be a structurally different population dynamics package, rather than an implementation within the same package using different datasets and or parameter values. It might be reasonable to consider multiple equally plausible implementations within the same package to be an ensemble, but we did not explore that here. Instead, we included both models that had been implemented for Atlantic herring by the lead assessment scientist: ASAP (Miller and Legault (2015)) and SAM (Nielsen and Berg (2014)); see the main body of the text and appendix B2 for descriptions. The models are similar in that they are age-structured single species population dynamics models. They differ in their approach to parameter estimation and treatment of uncertainty. Both models were developed iteratively by the lead assessment scientist to produce the desired level of fit diagnostics, although the ASAP model received further iterative development during the working group meeting while the SAM model did not.

Stock assessment models produce many estimated parameters and outputs. We determined quantities of interest for MMI using the ToRs for the assessment. Therefore, we focused on model derived quantities: spawning stock biomass (SSB), recruitment, fishing mortality (F), and reference points. Uncertainty in these quantities is of interest in some management frameworks as well. We did not attempt to apply MMI to projections and catch advice, for this example, but discuss possible approaches below.

In this example, we take a simple average of the derived estimates from the "best" model from each modelling framework. We calculate the confidence intervals using the minimum CI from the two models as the lower bound and the maximum CI as the upper bound at each point in the time-series. For these examples, approximate 80% confidence bounds are shown. This approach does not consider the uncertainty and information potential in all possible models or even all tested models, but is consistent with the objective to keep the ensemble process simple and achievable within an assessment timeframe.

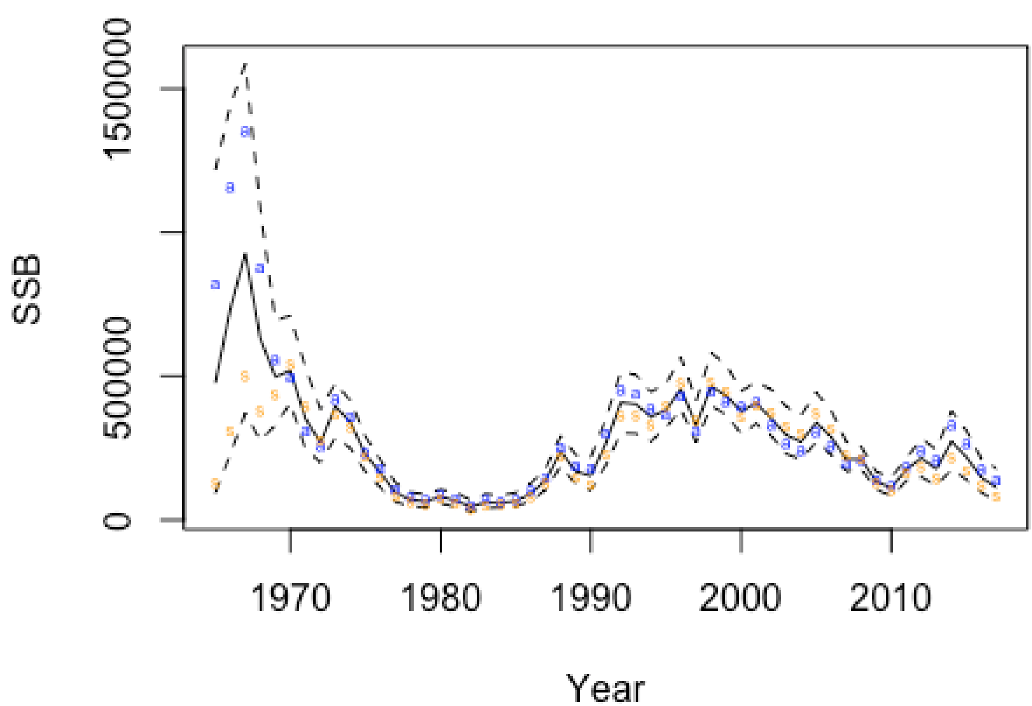
Results

Taking the simple average of output SSB, recruitment, and F from ASAP and SAM demonstrates where the models are similar and where they diverge. Population indices such as SSB and recruitment are fairly similar across both the herring and mackerel examples, with the model average not greatly different from individual models. Estimates of fishing mortality are shown, but we note that the selectivities estimated by the models are very different, such that comparing Fs is conceptually more difficult. Nevertheless, the two models estimate similar Fs for some portions of the time-series, and divergent Fs in others.

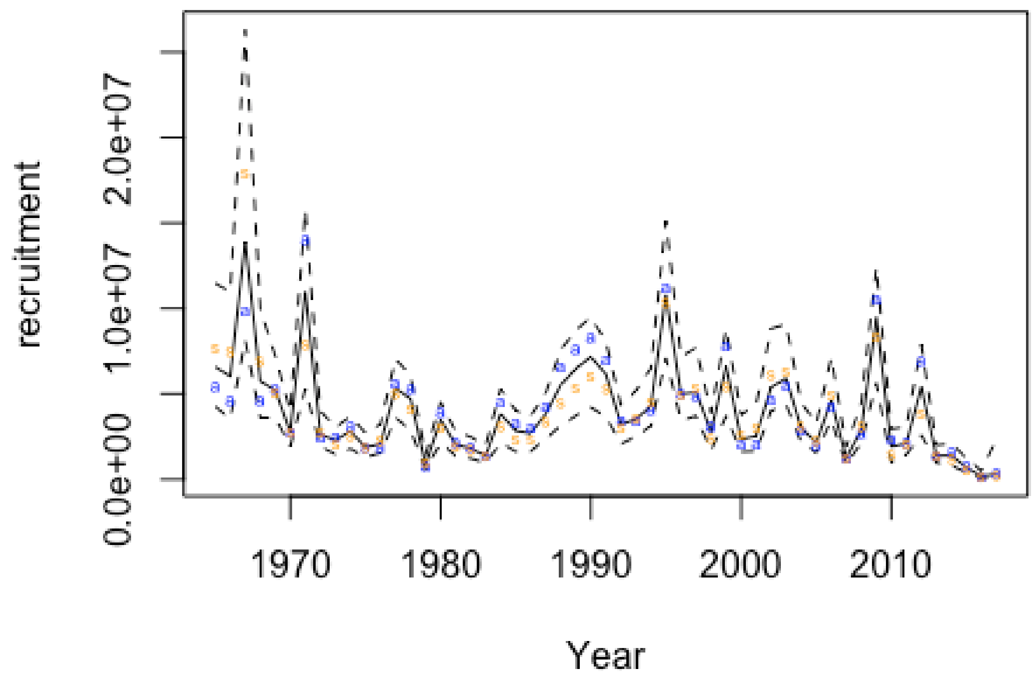
Perhaps more important than a measure of central tendency between models is the estimate of uncertainty when considering both. The "envelope" around the model average estimate contains both models by design, but still shows where our certainty is greater or lower across the estimation period.

Estimates of stock status were derived from the model outputs and as expected, the model average estimate falls directly between them. In this case, each model estimated different status, and the model average is right on the borderline with wide intervals.

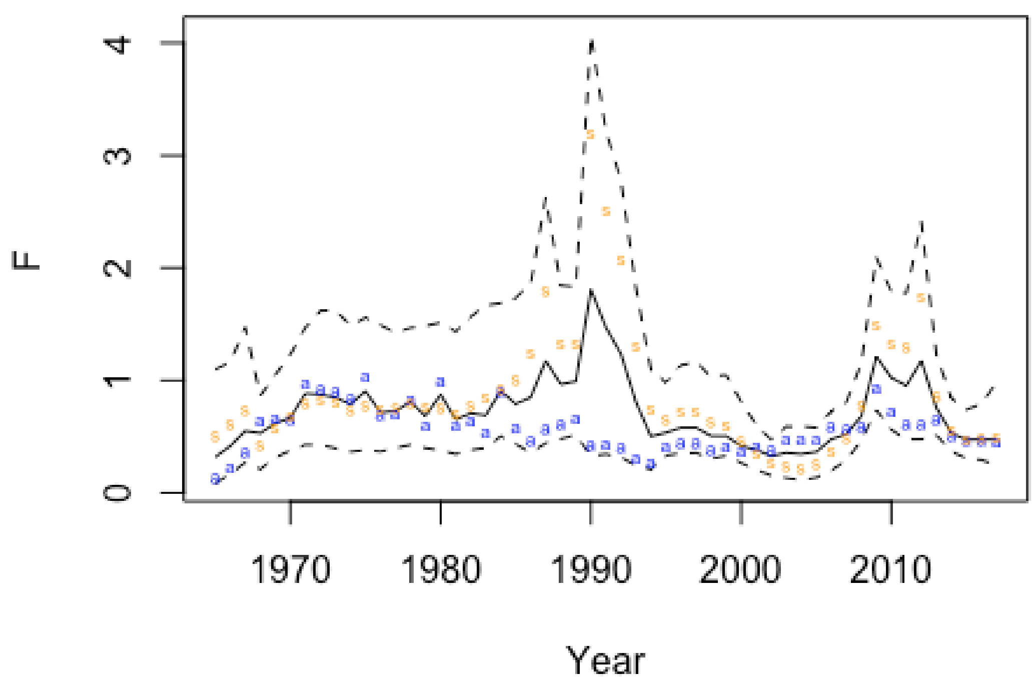
Burnham, K.P., and Anderson, D.R. 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. *In* 2nd editions. Springer-Verlag, New York. Available from //www.springer.com/us/book/9780387953649 [accessed 11 May 2018].



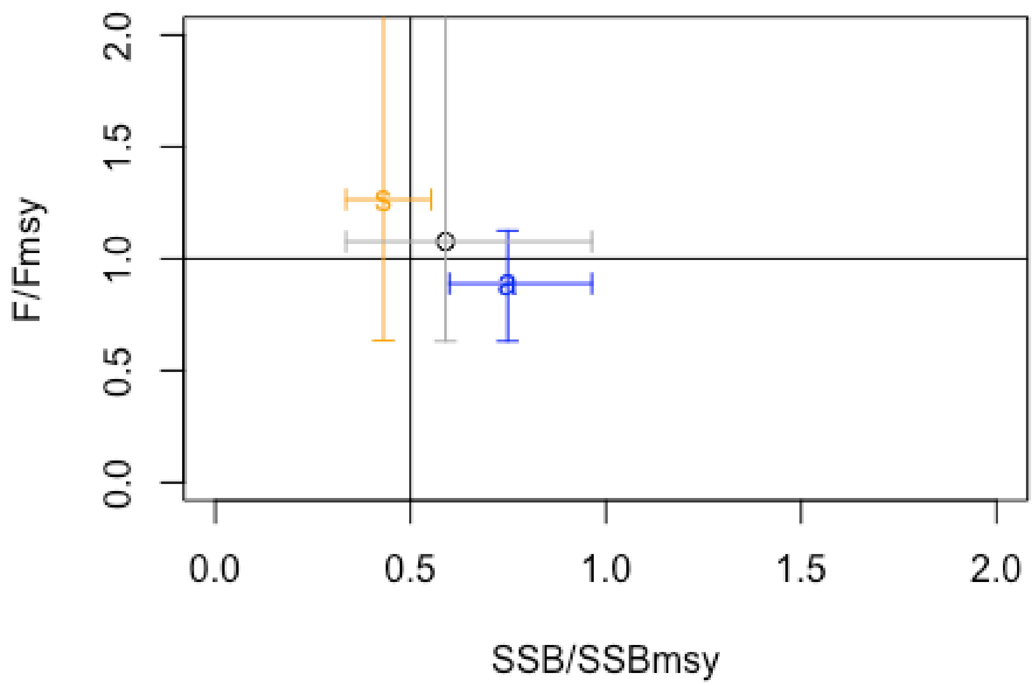
Average Herring SSB estimates with CIs



Average Herring recruitment estimates with CIs



Average Herring F estimates with CIs



Herring stock status from ASAP (a), SAM (s) and average

Exploring ecosystem effects of management

- Used a multi-model ensemble, with 4 multi-species models, and a Bayesian belief network to show that the Nash equilibrium is probably better than MSY.
- Proof of concept that shows that the ensemble model of Spence *et al.* (2018) can be used as an input to an external model.
- Current work is looking to combine single-species and multi-species models to make seamless predictions across a range of timescales.

Four ecosystem models were used to run three different fishing scenarios from 2013: FMSY, the maximum sustainable yield, Nash equilibrium (Thorpe *et al.*, 2016) and status quo. The landings were calculated for the most economically important species in the UK. A Bayesian belief network was used to estimate the economic value of landings for the UK. The multi-model ensemble concluded that probably the Nash equilibrium would be the best of the three future scenarios for the UK.

A current ongoing project at Cefas has sped up the fitting of the ensemble model under specific assumptions using a Kalman Filter. This would make the model more useful for the modelling community. Additionally, work is being done to exploit the predictive power of different models in time, e.g. formalising the idea that some models are good on the short term but others are good on the long-term. The aim being to combine single-

species and multi-species models to make seamless predictions across a range of time-scales

Spence MA, Blanchard JL, Rossberg AG, *et al.* A general framework for combining ecosystem models. *Fish Fish*. 2018;00:1–12. https://doi.org/10.1111/faf.12310

Thorpe RB, Dolder PJ, Reeves S, Robinson P, Jennings S; Assessing fishery and ecological consequences of alternate management options for multispecies fisheries, *ICES Journal of Marine Science*, Volume 73, Issue 6, 1 June 2016, Pages 1503–1512.

Comparing F values

Reported F levels aim to provide a univariate metric to reflect the fishing pressure of the stock as a whole. Standard ICES Fbar over a given age range may perform well on this, but it may not do. Different methods of computing summary F statistics were presented for a range of ICES stocks, looking at different age ranges and different weighting methods (unweighted, by catch number at age, by catch weight at age). There were no differences that were consistent across all stocks. In general stocks with stable selectivity and age structure were fairly insensitive to changes in methodology. However, some stocks where age structure or selectivity varies have obvious differences in the trends and features of these time-series. As a consequence, a poor choice of F summary statistic may misidentify trends in F, and thus not well reflect the overall fishing pressure over time. A series of comparisons showing the differences in both mean value and time trends in the F values accounting for these uncertainties was presented at WGSAM 2018, in order to gain feedback in the process of preparing a manuscript. Note that these differences and difficulties arising from a choice "F" statistic are in addition to those arising from the range of different methods of computing Fmsy values – F0.1, Fmax, Yield-per-recruit vs. spawners-per-recruit, reducing F to meet a precautionary criterion and so on. The choices involved in these calculations have the potential to result in different Fmsy values, making comparison between Fmsy values difficult.

Annex 8: ToR E. Test performance and sensitivity of ecosystem indicators

A simulation study of trend detection methods for IEA (contributed by Sean Hardison Sean.Hardison@noaa.gov)

Preliminary results were presented of Monte-Carlo simulations aimed at assessing the ability to identify statistically significant trends from time-series of varying length and autocorrelation regimes using a range of methods. The work is aimed at understanding current limitations with respect to the development of management advice using indicators of varying lengths, as well as our ability to comment on "recent" trends, which managers have a particular interest in. A manuscript detailing the results and their implications is currently in draft by Sean Hardison, Charles Perretti, and Geret DePiper, and Andy Beet.

The research simulates time-series with characteristics similar to the current suite of indicators utilized by the Northeast U.S. Integrated Ecosystem Assessment program, and compares three statistical approaches in their ability to correctly identify trends in the data. The analysis currently includes 1000 simulations of the combination of four trend (no, weak, medium, and strong trends), three autocorrelation (no, medium, and strong autocorrelation), and 3 time-series length (10, 20, and 30 years) regimes. Trends themselves were assessed using three statistical tests (Mann-Kendall, Mann-Kendall with prewhitening, and Generalized Least Squares, the latter of which is actually a set of four models, with the best fit selected based on AICc).

Results indicate that our ability to identify recent trends in the data is limited, and statistical approaches were biased by even moderate amounts of autocorrelation at small samples sizes ($N < 30$).

After the presentation, Sigrid Lehuta gave an overview of the Sept 28 2018 WKINTRA workshop on ecosystem indicator analysis (where the above work was also presented). For analyzing multiple ecosystem indicators, PCA has been popular in the past. A recent paper shows that PCA is not robust to autocorrelations, so other methods should be explored instead. This was the first meeting of WKINTRA, and will be at least 2 other meetings (after next ICES conference, and the following January). Methods will be similar to the above, applying different analytical tools to simulated time-series with known qualities to examine the strengths and weaknesses of different methods.

Annex 9: ToR F. Metanalysis of impact of top predators on fish stocks in ICES waters

Consumption estimates in the New England Atlantic herring assessment (Sarah Gaichas)

In the 2018 benchmark herring assessment, consumption by groundfish was estimated from food habits data. In the 2012 benchmark assessment, this type of information was used to justify a step increase in M starting in the early 1990s, because the "consumption" in tons implied by the assessment M matched this time-series reasonably well. This step change in M also happened to correct a retrospective pattern in the 2012 benchmark assessment. However, by the 2015 update assessment the "consumption" implied by the increased M had diverged from the consumption estimated from groundfish food habits data. Further, the retrospective pattern returned. Finally, in 2018 a smoothed function of the predation mortality implied by consumption appeared generally flat (Figure 1), so while consumption was estimated (Figure 2), it was not used in the assessment, and M was assumed constant over age and time.

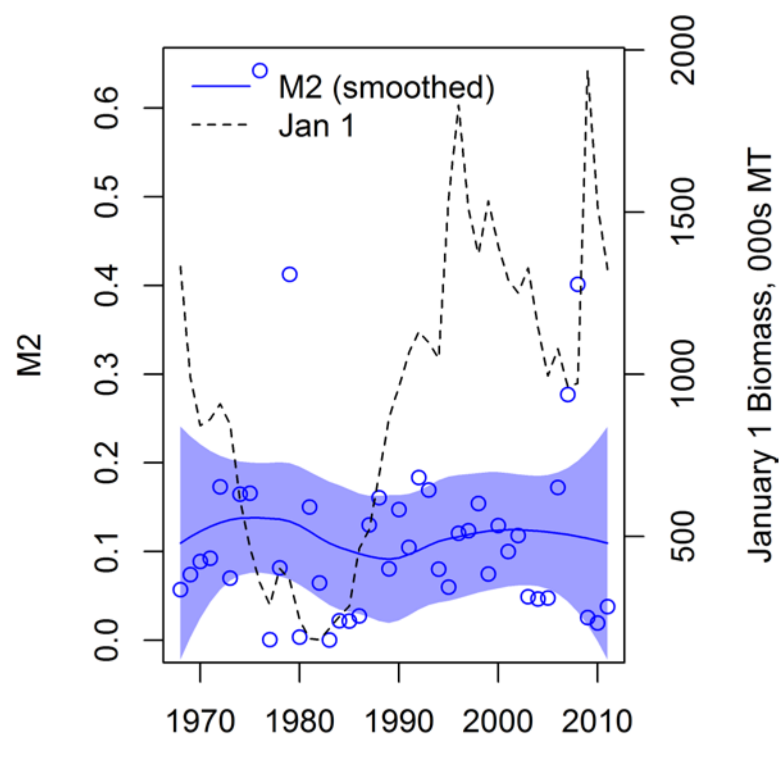


Figure 1. Proxy estimate of natural mortality due to predation (M2) and January 1 biomass of herring, 1968–2011. M2 smoother is loess with span = 0.8 and 95% ci.

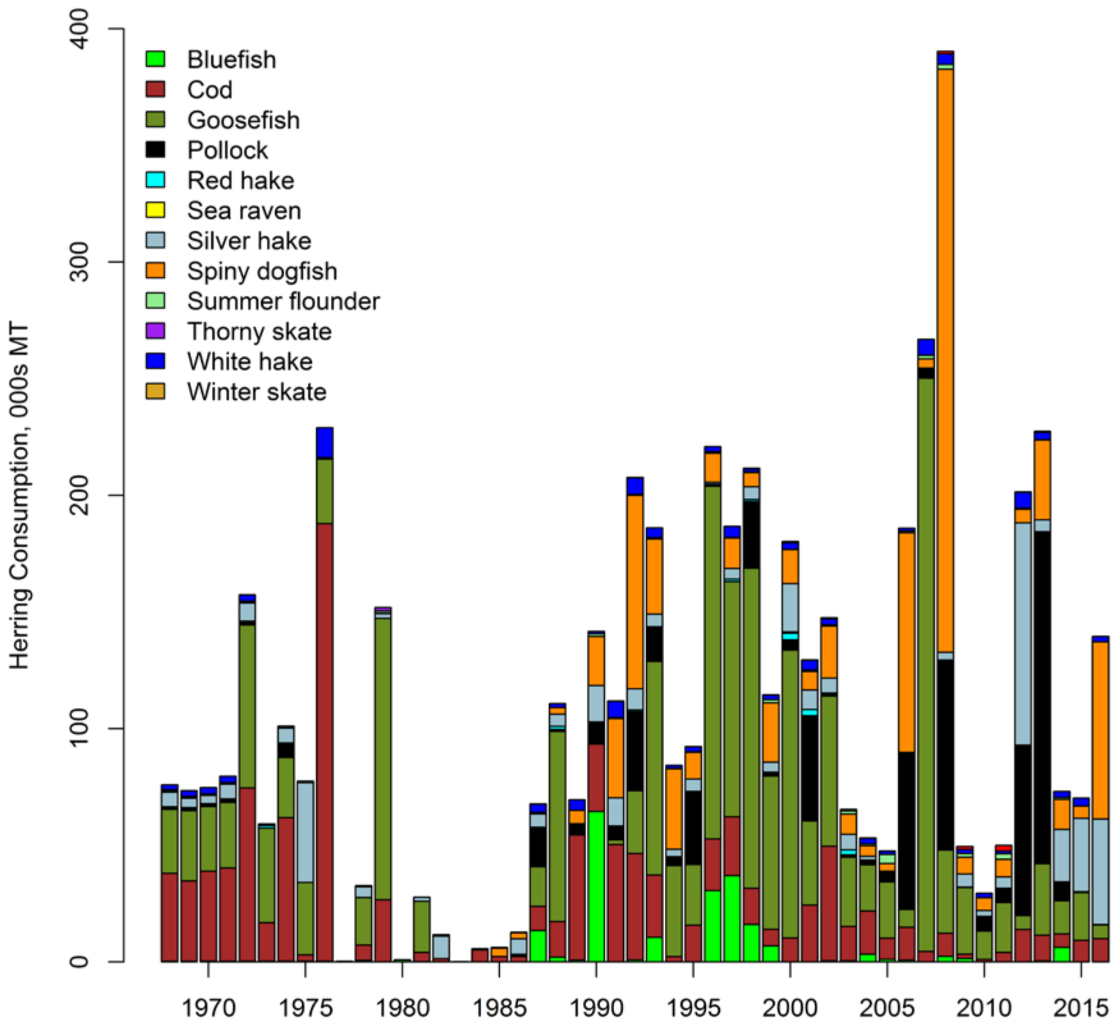


Figure 2. Time-series of herring consumption (000s MT) by 12 fish predators.

Annex 10: ToR G. Explore the consequence of multispecies, mixed fisheries interactions and environmental factors in practical multi-species advice for fisheries management

Multispecies and mixed fisheries management strategy evaluation in the North Sea

Robert Thorpe gave a presentation on a management strategy evaluation (MSE) being conducted using a length-structured multispecies and mixed fisheries model of the North Sea fish community. Five candidates for a community MSY (CMSY), the 21-stock Nash equilibrium, one based on single species assessments, and ones based upon the top, middle, and bottom of the ICES "pretty good" yield ranges were evaluated using a variety of Harvest Control Rules (HCR), with outcomes being assessed in terms of average risk of stock depletion and gross revenue (price x catch). The MSE was carried out with an ensemble of 63 models with stochastic recruitment, repeated 100 times for each scenario. In the absence of an HCR, we find that the lower PGY ranges are the safest option and the Nash equilibrium the highest yielding, with the other options being sub-optimal. Application of an HCR cuts risk and reward, the former more than the latter such that the HCR is useful. The impact of the HCR depends on its functional form and the point at which yield is reduced (MSY Btrigger). We find that the optimum choice for CMSY depends on societal views of acceptable risk, with no clearly optimum solution. However, the upper part of the PGY ranges is never a good choice.

Previous work with LeMans (Thorpe 2015, 2016, 2017) has focussed on constant harvesting strategies – fishing at the same mortality regardless of stock status to evaluate the long-term impacts of the strategy. But in practice, fishing would not take place in this way, because it would be reduced if the stock status is poor. This is often done via a harvest control rule (HCR), a pre-agreed management procedure which determines how the fishing mortality target varies with stock status. In this study we ask what is the best way of achieving multispecies MSY (assuming this is defined as achieving the maximum possible yield for an acceptable risk using HCRs within a management strategy evaluation framework. Thus, we take account of model parameter uncertainty, management target uncertainty, and fleet management uncertainty as well as management implementation uncertainty, evaluating outcomes in terms of risk and reward. A schematic of the experiment design is shown in Figure 1.

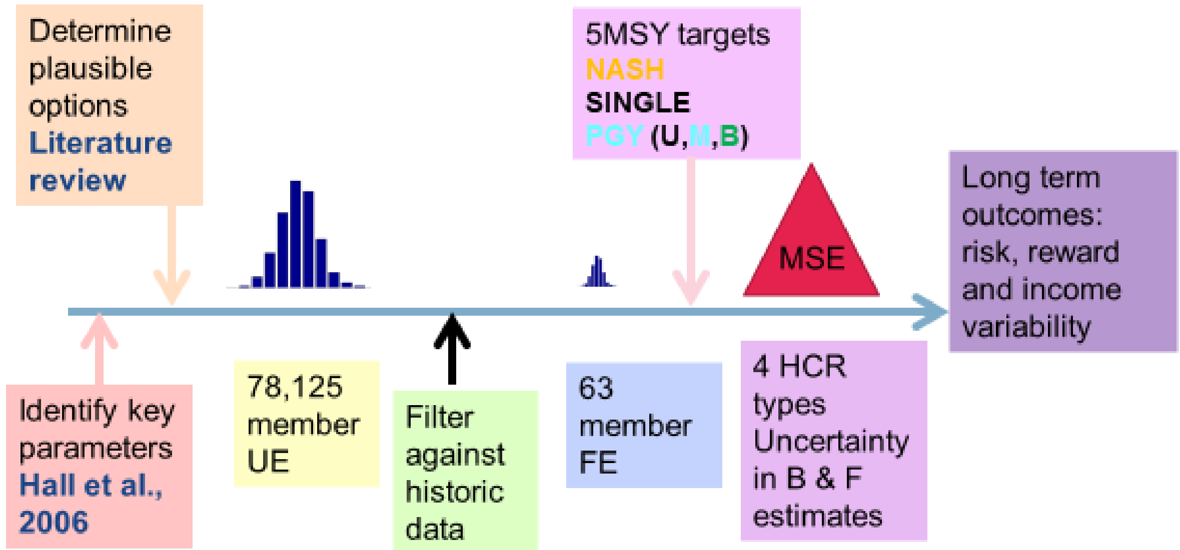


Figure 1. Schematic of the experimental design for the management strategy evaluation.

We use the same 63 member ensemble as in Thorpe *et al.* (2017), with each ensemble being evaluated 100 times to take account of stochastic variation in recruitment. 5 candidates for a community MSY were considered, one based on 2012 single species assessments (Thorpe *et al.* 2015), one based on the Nash equilibrium (Thorpe *et al.* 2017) and three based on the bottom, middle, and top of the "pretty good yield" ranges as defined by ICES. Of the 21 stocks, 7 (cod, haddock, whiting, sole, plaice, herring, and saithe) have published ranges as in Table 1.

Table 1: ICES estimates of "pretty good yield" for 7 North Sea stocks.

Stock	F-PGY upper	F-PGY central	F-PGY lower
Herring	0.39	0.33	0.24
Sole	0.37	0.20	0.113
Whiting	0.15	0.14	0.14
Plaice	0.30	0.21	0.146
Haddock	0.194	0.194	0.167
Cod	0.46	0.31	0.198
Saithe	0.49	0.36	0.21

The other 14 PGY ranges were generated using the following assumptions. 1) The average F across all stocks was maximised, 2) no fleet can have more than three times the effort of another, when they are both expressed as effort relative to the average effort between 1990 and 2010, and 3) there are no discards, so the fishery is limited as soon as the first choke limit is reached. Figure 2 shows the Fs that result from these assumptions.

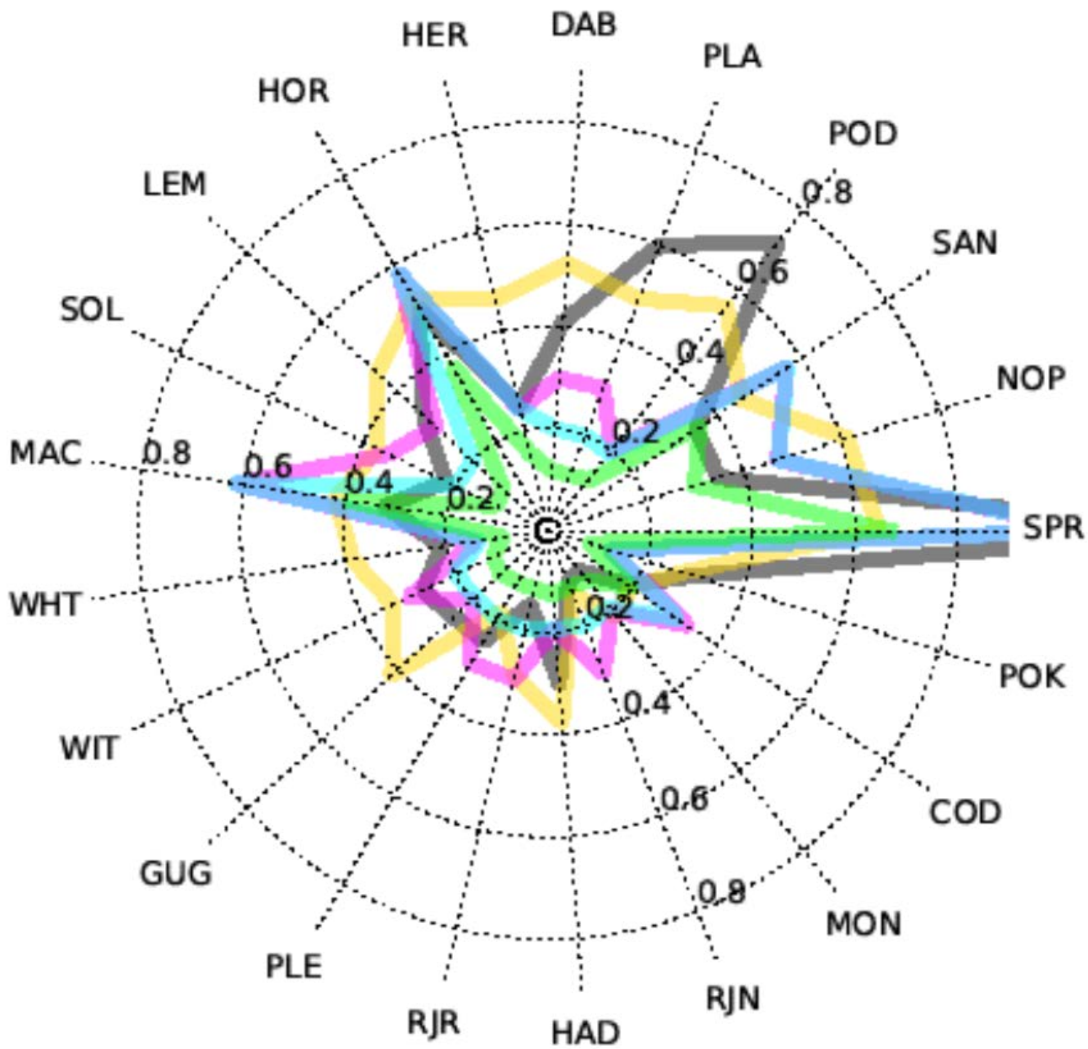


Figure 2. Radar plot of community MSY for a) estimates based on single species assessments (black), b) 21-stock stochastic Nash equilibrium (gold), c) upper PGY ranges (magenta), d) mid PGY ranges (cyan), and e) lower PGY ranges (green).

Our management strategy evaluation considered four types of harvest control rules, which are shown in Figure 3.

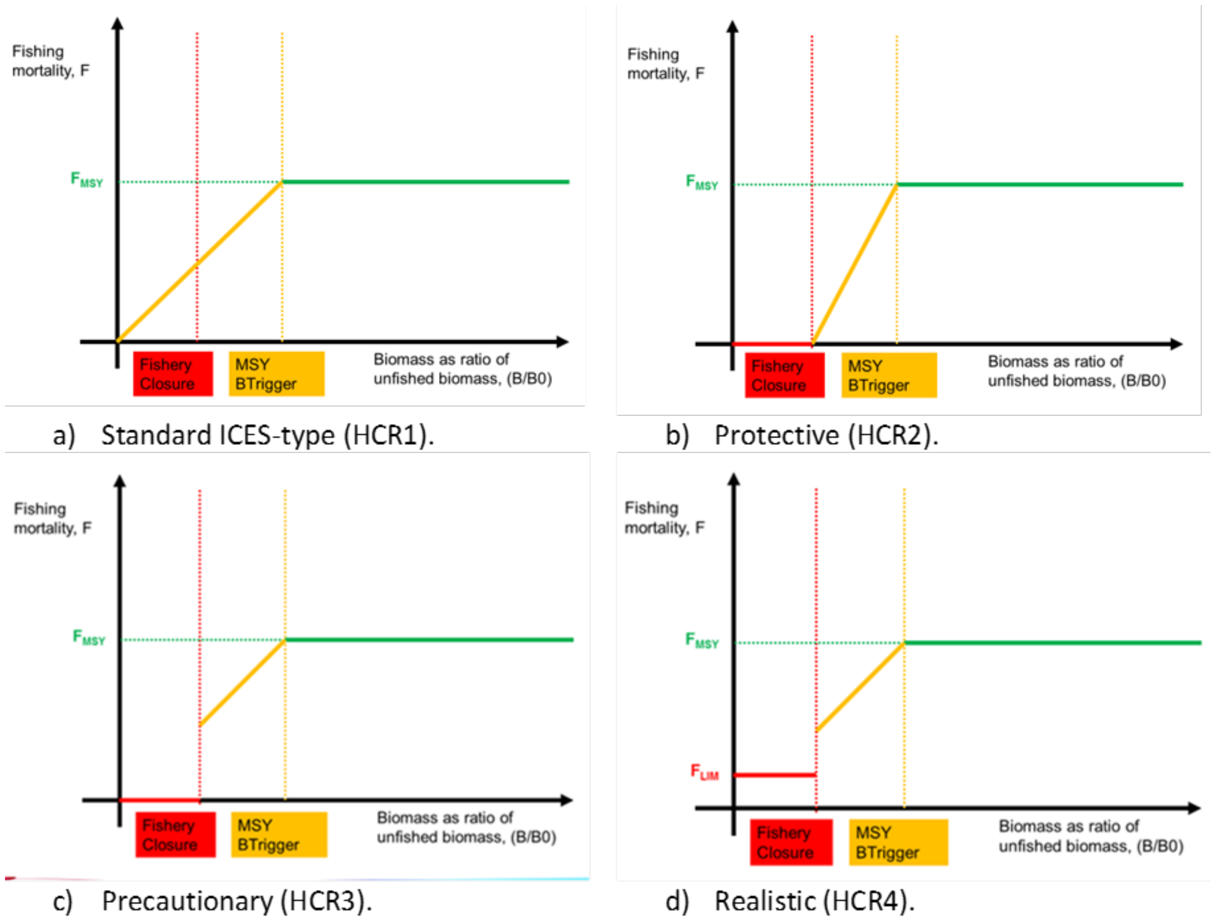


Figure 3. Schematic of the 4 types of harvest control rules used in the management strategy evaluation.

Within the MSE, the ensemble model acts as the operating model. Stock status is assessed by taking the ensemble mean biomasses adjusted by a log-normal error term of given size from 0 to 50%. The harvest control rule is then used to generate a target F for the stock, which is implemented with lognormal uncertainty of given size from 0 to 30%. Stock status is assessed annually, after which the newly ascertained F is applied for the next year.

For the reference case with constant F and no HCR, results are shown in Figure 4.

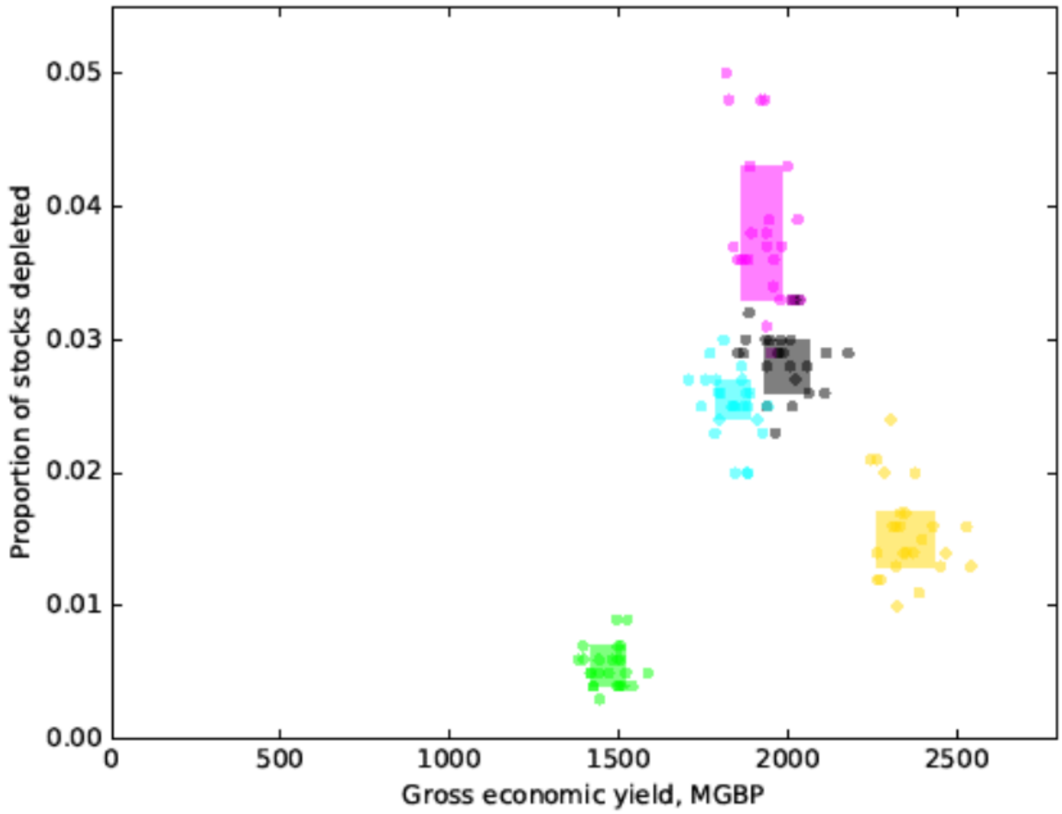


Figure 4. Risk – reward outcomes for constant F strategies. Black = single species, Gold = Nash, GreEn = L-PGY, Cyan = M-PGY, Magenta = U-PGY.

We find that the Nash equilibrium gives the highest yield, and L-PGY is the safest. A choice between these would depend on societal risk appetite. The other solutions are sub-optimal. Application of an HCR reduces both risk and yield, the former more than the latter. The nature of the reduction depends upon the form of the HCR, the choice of MSY Btrigger, and the level at which a stock is considered depleted.

Overall we find the following:

a) The best outcome depends upon societal views of risk and reward – there is no choice of CMSY that is clearly optimal, although some can be dismissed as sub-optimal.
b) The upper PGY ranges are never a good choice (consistent with Thorpe *et al.* 2017).
c) The annual operation of an HCR reduces both risk and yield. Yield reductions are more modest than those of risk, making the HCR a valuable management tool, independent of its form (amongst those considered here), definition of risk, or definition of MSY Btrigger.

References:

S.J. Hall, J.S. Collie, D.E. Duplisea, S. Jennings, M. Bravington, and J. Link (2006) A length-based multispecies model for evaluating community responses to fishing. Canadian Journal of Fisheries and Aquatic Sciences, 63, 1344–1359.

R.B. Thorpe, W.J.F. Le Quesne, F. Luxford, J.S. Collie and S. Jennings (2015) Evaluation and management implications of uncertainty in a multispecies size-structured model of population and community responses to fishing, Methods in Ecology and Evolution, 6, 49–58. doi: 10.1111/2041–210X.12292

R.B. Thorpe, P.J. Dolder, S. Reeves, P. Robinson, and S. Jennings (2016) Assessing fishery and ecological consequences of alternate management options for multispecies fisheries, ICES Journal of Marine Science, 2016, DOI 10.1093/icesjms/fsw028

R.B. Thorpe, S. Jennings, P.J. Dolder (2017) Risks and benefits of catching pretty good yield in multispecies mixed fisheries, ICES Journal of Marine Science, DOI 10.1093/icesjms/fsx062

An Ecosystem-Based Management Procedure for Multispecies Fisheries on Georges Bank (contributed by Mike Fogarty)

Georges Bank is widely recognized as a highly productive marine ecosystem. It has supported generations of fishing communities on the northeast seaboard since the early 18th century when offshore fisheries first developed in the United States. The Georges Bank ecosystem was subject to a massive impact with the arrival of distant water fleets in 1961, resulting in the decimation of a number of fish stocks in a pattern of sequential depletion (Fogarty and Murawski 1998). The history of groundfish management on Georges Bank since then has involved seemingly intractable problems related to the pervasive technical and biological interactions in this system. In the following, we describe elements of an Operational Management Procedure (OPM) for multispecies fisheries designed to address these challenges. We focus on a system approach centered on the concept of functional group management. For our purposes these functional groups comprise species that are caught together and share similar life history characteristics and trophic positions. They lie at the intersection of fishery-related and ecological interactions.

OPMs are designed to establish a setting in which (potentially) simple management rules are identified and rigorously tested to address objectives for management developed in a transparent process with stakeholder involvement. At the request of the New England Fisheries Management Council (NEFMC), options for Ecosystem-Based Fisheries Management (EBFM) are being explored by its EBFM Plan Development Team, including the work described below.

The main elements of the multispecies OPM under consideration involve (1) the establishment of a dynamic ceiling or cap for total fishery removals from the Georges Bank ecosystem (2) specification of catch allocations to defined Fishery Functional Groups (FFGs). The sum of these catch allocations by FFG cannot exceed the system ceiling. (3) identification of floors or thresholds below which individual species cannot be driven without invoking remedial action. This Floors and Ceilings approach is now being tested by simulation to assess its performance using a size-structured multispecies multifleet model Hydra (Gaichas *et al.* 2016). A flow diagram of the principal elements of the model is provided in Figure 1.

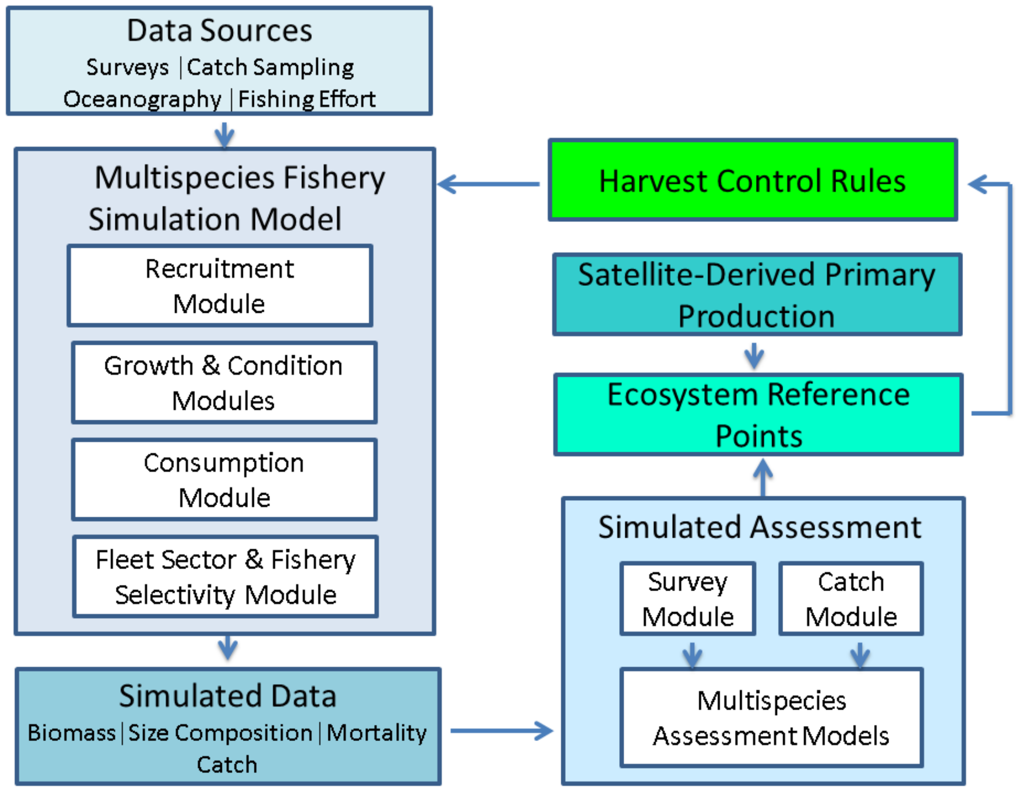


Figure 1. Components of the simulation model used to test management procedures in Hydra.

Hydra focuses on a 10 species subset of the Georges Bank fish community: Atlantic cod (Gadus morhua), haddock (*Melanogrammus aeglefinus*), silver hake (*Merluccius bilinearis*), winter flounder (*Pseudopleuronectes americanus*), yellowtail flounder (*Limanda ferruginea*), monkfish (*Lophius americanus*), spiny dogfish (*Squalus acanthias*), winter skate (*Leucoraja ocellata*), Atlantic herring (*Clupea harengus*), and Atlantic mackerel (*Scomber scombrus*). These species account for a major fraction of the total landings of fish species managed by the New England Fishery Management Council.

The harvest control rules examined here determine overfishing at the FFG level but overfished (or depleted) status at the species complex or individual species levels. As a prelude for undertaking a full Management Strategy Evaluation to be initiated in 2019, we undertook preliminary exploration of 6 principal scenarios with four levels of exploitation nested within each (Table 1)

Table 1. Scenarios Tested in simulation studies of the EBMP.

Scenario 1 Threshold exploitation (no ramp down) at Ex=0.15, 0.2, 0.25, 0.3 and Floor=0.2 of unfished biomass applied at the species complex level
Scenario 2 Threshold exploitation (no ramp down) at Ex= 0.15, 0.2, 0.25, 0.3 and Floor=0.2 of unfished biomass applied at the individual species level
Scenario 3 Threshold exploitation (no ramp down) at Ex= 0.15, 0.2, 0.25, 0.3 and

Floor=0.2 of unfished biomass for each species except vulnerable species (winter skate and dogfish) with a Floor=0.3 of unfished biomass) applied at the individual species level
Scenario 4 Ramp-down exploitation using 'steps' at Ex=0.15, 0.2, 0.25, 0.3 and Starting at B/Bo = 0.4 applied at the species complex level
Scenario 5 Ramp-down exploitation using 'steps' at Ex=0.15, 0.2, 0.25, 0.3 and Starting at B/Bo = 0.4 applied at the individual species level
Scenario 6 Ramp-down exploitation using 'steps' at Ex=0.15, 0.2, 0.25, 0.3 and Starting at B/Bo = 0.5 applied at the individual species level for vulnerable species (winter skate and dogfish)

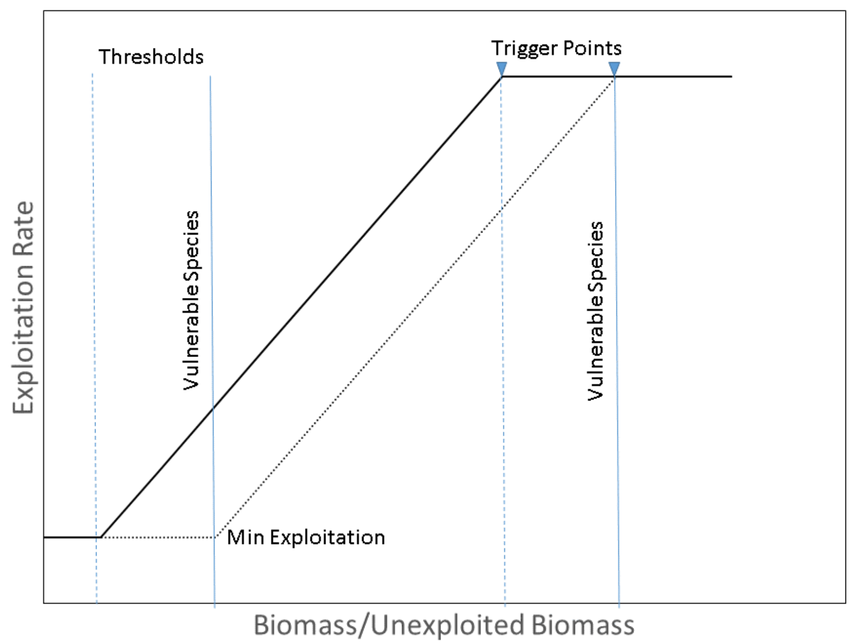


Figure 2. Structure of the ecosystem-based harvest control rules tested. Overfishing is determined at the species complex level. Overfished status is determined at the species complex or individual species levels (see details in Table 1).

Performance metrics

To evaluate fishery performance, we examine Catch, Biomass, and the fraction of simulation runs in which the species and/or functional group constraint (floors) was exceeded. We used the median result of the 500 member ensemble to compare different control rules and their variants but show the full range of results characterizing uncertainty with a focus on the interquartile range. In the simulations, we also output the associated revenues, the size composition of the catch and the population for each species. Additional metrics including measures of biodiversity are also part of the output.

The preliminary results indicate that:

Performance of fixed exploitation rate strategy was significantly worse for all metrics than ramp-down strategies at all exploitation levels

At exploitation rates as low as 0.15, performance of ramp-down strategies at the functional group and individual species levels, and the enhanced protection strategy for vulnerable species are very similar for all metrics.

At higher exploitation rates, the species-level and enhanced protection level strategy increasingly out-perform protections placed at the functional group level.

At highest exploitation rate examined (0.30), the enhanced protection strategy for vulnerable species pays the highest dividends.

Collectively, these simulation results suggest that defining overfishing at the species complex level and affording a biomass floor at the species level can sharply reduce the incidence of overfished status determinations.

Expanding beyond indicators: the MareFrame decision support framework for an EBFM of the Baltic fisheries (contributed by Valerio Bartolino *et al.*)

A generic framework for supporting decision within the context of ecosystem based fisheries was developed in the MareFrame project (Fig. 1) and presented to WGSAM. The framework is developed around the following main steps: (1–2) initial scoping process to frame the objectives and formulate potential alternative management strategies, (3) implementation of ecosystem models and scenarios according to the alternative management strategies, (4) identification of relevant indicators (i.e., utilities) to characterise the consequences of the alternative management strategies, (5) elicitation of stakeholders' preferences and weights on the main utilities of the system, (6) evaluation of trade-offs and ranking of the alternative management strategies within a decision support tool which brings ecosystem models output and stakeholders' preferences. All the steps of the framework (with the exception of step 3) involve a high level of interaction with stakeholders that takes the form of a co-creation of management solutions.

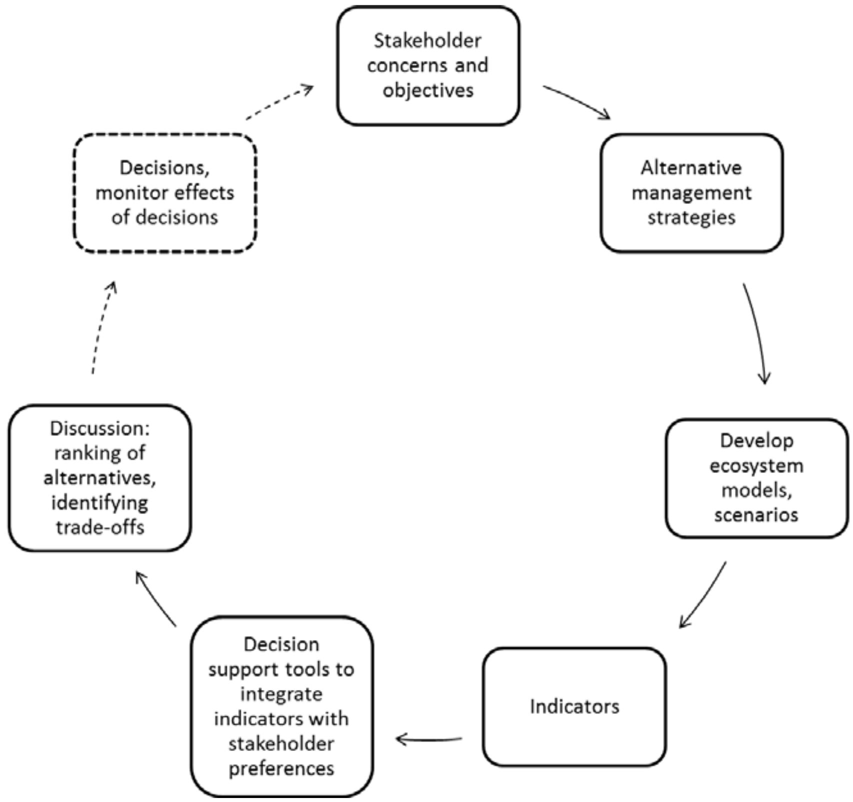


Figure 1. Generic Decision Support Framework for EBFM developed and tested in the central Baltic Sea.

The framework, which is generic in its essence, was tailored and tested in the central Baltic Sea case study of the MareFrame project, for the long-term sustainable management of the cod, herring and sprat fisheries considering trophic interactions among these three stocks, the uncertainty associated to different nutrient scenarios, alternative growth rates of the grey seal population, and socio-economic benefits for the fisheries. Testing of the framework was possible thanks to the active participation of the case study stakeholders which were representative of different profiles including managers, advisory groups, NGOs, fishing industry.

Three complementary ecosystem models – i.e. Ecopath with Ecosym (EwE), Gadget and a multispecies production model (MSPM) – were applied for this purpose. The high level of complementarity (Fig. 2) of the three models allowed inference on potential trajectories of the system at a population, community and ecosystem level.

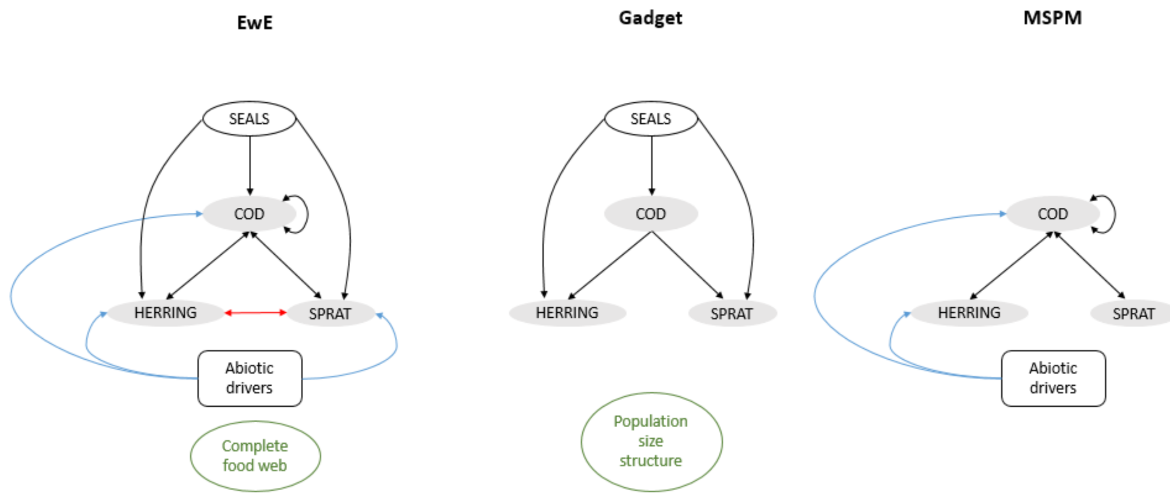


Figure 2. Model components and interactions in EwE, Gadget and MSPM models of the central Baltic Sea ecosystem. Black arrows represent predation, the red arrow competition, and blue arrows abiotic environmental effects. Single-headed arrows represent one-way interactions and double-headed arrows dynamic feedbacks. Green ovals denote components unique to a model.

The study evaluated the ecological, social and economic implications of the following alternative management strategies at medium- (2020) and long-term (2030):

- BAU (Business as usual) - average fishing levels estimated in 2011–2013
- MEY_{cod} (Maximum Economic Yield for cod) - maximize cumulative discounted profit of bottom trawlers and gillnetters derived from their cod catches
- MEY_{pel} (Maximum Economic Yield for pelagic stocks) - maximize cumulative discounted profit of the pelagic fisheries derived from their sprat and herring catches
- MEY_{all} (Maximum Economic Yield for all stocks) - maximize cumulative discounted profit of all the fisheries
- ES (Environmental Strategy) - maximize cod biomass compared to clupeids while maintaining both herring and sprat at viable levels

A set of 22 indicators were calculated from the three models outputs. These indicators describe performance of the alternative management strategies in achieving the goals and avoiding the concerns identified by the stakeholders. These indicators inform about the expected status in the fisheries socio-economic system, fish stocks and standard fishery related quantities such as fishing mortality rate, and other natural ecosystem components, and relate to some of the main MSFD descriptors (D1 and D3). Model outputs were used to inform a decision support tool (DST) based on Bayesian Influence Diagrams https://mareframe.github.io/dsf/dev/BBN2/DST.html?bbn=true&model=baltic

Preferences on the different levels of the 22 indicators (pre-binned for this purpose) were elicited from the stakeholders involved in the project.

In conclusion of the case study, the framework and its tools (ecosystem models and DST) were positively evaluated by the participants as it offers a way to:

- have a structured comparison and evaluation of alternative management strategies
- reduce ecosystem models complexity for their potential use in the advisory process
- synthesis information to non-scientists
- facilitate communication about choices and reasons
- improve commitment to the decisions
- understand trade-offs and conflicts
- improve potentially rigor and transparency in decision making
- Caveats of the framework and potential limits in upscaling its implementation to a real setting (outside the boundaries of a research project case study) were identified in:
- the need of multidisciplinary expertise
- high level and long-term individual engagement of both scientists and stakeholders
- difficulty to manage transparency if many stakeholders should be involved in a real setting
- high uncertainty in quantifying some of the key processes driving the central Baltic ecosystem
- unclear which body could have the authority to gather relevant stakeholders and should lead the process structured by this framework

Ecosystem Fmsy Project

The Ecosystem Fmsy project has been conducted to attempt to find practical steps for improving the current Fmsy/Ftarget fishing levels by including more ecosystem realism in a way that could enter management in the short term. Attempting to produce full multispecies advice has not, in general, been successful in entering tactical management. Statistical work in this project indicates that between 1/3 and 1/2 of ICES stocks exhibit density dependence in growth, but such density dependence is only considered in a very small number of managed stocks within ICES. We hypothesize that the requirement to confront trade-offs in multispecies systems is a large part of the reason for this lack of take up in management. Hence to project seeks to simply improve the existing single species Fmsy targets include some partial ecosystem information. Production (and Ecopath) models were run for a large number of ICES stocks in order to include density dependence and hence implicitly ecosystem considerations, into target fishing levels. These production models produced alternate candidate “multispecies” Fmsy values, which are mostly higher than the existing single species values. However, it is important to not simply accept these values uncritically. A first step is to compare the results with known biological and stock knowledge to identify if these values are reasonable for the current state of the stock (for example stocks that do not violate production model assumptions that catch drives stock, that do not show strong regime shifts over the tuning series, and where the stock is either high or rising at current fishing levels and could thus sustain higher F levels). A further consideration is that density dependent effects are likely to be strongest at high stock sizes, and in fact the one ICES stock (NEA cod) where a "two step" HCR accounts for density dependence/carrying capacity is one which is at high stock size. The project also highlighted MSE tools that are able to account for these density dependent processes, and that the simple Fmsy values can thus be converted into precautionary Ftarget values.

WGSAM supports this effort to improve the current management targets in this way. Although group considers that it is clearly not precautionary to simply adopt the proposed Fmsy values for a range of stocks, the group does consider that using these values to identify a subset of stocks to conduct further MSE-style evaluations with a view to potentially revising existing target reference points represents a viable approach to improving ICES management.

Multispecies MSE in the NAFO area: Flemish Cap case study

As introduced in Annex Tor A, the EU has promoted the project SC05 *"Multispecies Fisheries Assessment for NAFO"* in support of the development of the NAFO EAF roadmap. The Flemish Cap cod, redfish and shrimp fishing system has been modelled within a multispecies gadget model, GadCap (Pérez-Rodríguez *et al.* 2017). A main goal of the SC05 project is contributing to the development of the NAFO roadmap for an EAF exploring alternatives to incorporate the multispecies approach into the fisheries advice process. Specifically, the multispecies model GadCap has been used in first place to provide estimates of predation mortality that were used during the 3M cod benchmark exercise (Pérez-Rodríguez and González-Costas 2018). A more advanced approach has been the development of a multispecies MSE framework (msMSE) integrating the multispecies model GadCap as operating model within an a4a-MSE framework (Jardim *et al.* 2017). GadCap provides information about the "real" stocks, survey and commercial fleets that, once modified by the observation error model, is used for stock assessment in the management procedure module. Within the framework each of the three stocks has its own independent management procedure module. The current settings allow for a shortcut assessment, with and without assessment error, but also an assessment using an a4a SCAA model, that can also consider errors in the observation of survey and commercial information.

This msMSE framework has been used to design and test HCRs which reference points have been estimated following the NAFO standard protocols for single species approach, and compare its performance with HCRs defined considering the interdependent productivity of the three stocks, i.e. from a multispecies approach. Long-term simulations were run considering multiple combinations of Fs for cod, redfish and shrimp. The results show the influence that variable fishing strategies on predators (cod and redfish) would have on the prey stocks (shrimp and redfish). It is especially evident the impact that different fishing strategies on cod would have in the productivity of shrimp and redfish. In the case of shrimp, only when very high or very low fishing pressure on cod is implemented, the shrimp SSB reaches values above Blim (Figure 3). This pattern is due to the importance of cod as predator of redfish and shrimp, and the relevance of redfish as predator of shrimp.

The risk assessment (considering recruitment uncertainty) of the different one stage hockey stick HCRs indicates that the single sp oriented HCRs are not precautionary for cod and shrimp. Likewise, due to the strong trophic interactions between cod, redfish and shrimp, if shrimp wants to be maintained above Blim, fishing pressure on cod and redfish has to be so high that when the recruitment uncertainty is considered, the risk of being bellow Blim for this stocks in very high. In conclusion, no any combination of multispecies HCRs would maintain the SSB of the three stocks above Blim at the same time. This result indicate that multispecies HCRs have to be designed disregarding one or two of the other species in the system. For example, in the case that shrimp is disregarded, (i.e., the impact of a given fishing strategy on shrimp stock is disregarded), a number of combinations of Fs (HCRs) is obtained for which the risk of being below Blim at the same time for cod and redfish. Additionally, as an exploratory exercise, a two stage hockey sticks HCR for cod was simulated, with the intention of testing if reducing an excessive predation capacity from cod. These two stages HCRs were designed with an increase in fishing pressure on cod to F=0.55 when the SSB was above 45 000 t. This two stage HCR clearly reduced the risk of being below Blim bot for cod and redfish.

This study allowed concluding that:

- Single species F reference points were not precautionary for cod and shrimp.
- The results suggest that it is not possible having the 3 sps above Blim
- Disregarding one stock (shrimp or another stock) may allow finding precautionary multispecies reference points for the others.
- Considering assessment error in the shortcut option reduces slightly the risk due to underestimate of population size.
- The results suggest that the two stages HCRs for cod reduces predation and increases probability of cod, redfish and somehow shrimp being above Blim.

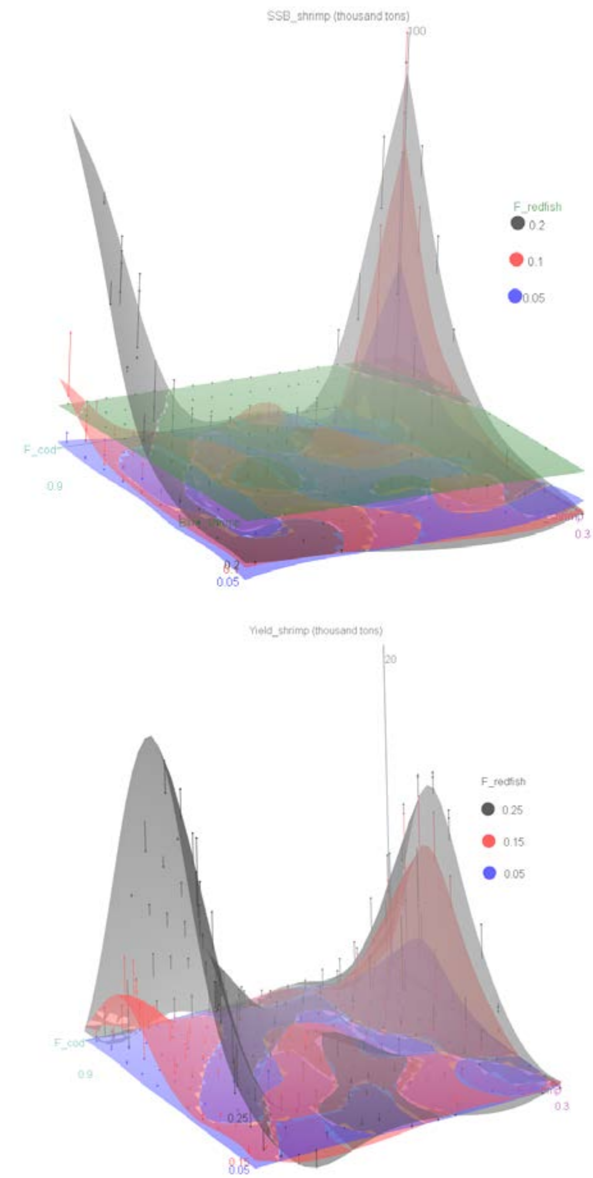


Figure 3.- Estimated shrimp SSB (left panel) and yield (right panel) using the updated Flemish Cap multispecies gadget model GadCap shrimp when different F values are applied on cod (Z axis) and shrimp (X axis) for three three different Fishing pressure levels on redfish.

Pérez-Rodríguez, A., and González-Costas, F. 2018. Estimates of natural predation and residual mortality for the Flemish Cap cod NAFO SCR Doc. 18/025.

Jardim, E., Scott, F., Mosqueira Sanchez, I., Citores, L., Devine, J., Fischer, S., Ibaibarriaga, L., Mannini, A., Millar, C., Miller, D., Minto, C., De Oliveira, J., Osio, G., Urtizberea, A., Vasilakopoulos, P. and Kell, L., Assessment for All initiative(a4a) - Workshop on development of MSE algorithms with R/FLR/a4a, EUR 28705 EN, Publications Office of the European Union, Luxembourg, 2017, ISBN 978–92–79–71290–6, doi:10.2760/18924, JRC106750

Exploring multi-species MSY in a mixed fishery: accounting for technical interactions among species

The single species approach for managing at FMSY assumes that fisheries can be managed using F values and does not account for changes in catchabilities as the stock size changes. It makes the approach particularly difficult to implement in mixed fisheries where species are caught simultaneously because the effort required to reach one species' Fmsy might be very different from the one required for another species. In addition, fishers change their strategies (allocation of effort on métiers) to adapt to stock availability and markets while Fmsy computations assume constant and often linear relationship between effort and fishing mortality (constant selectivity and catchability. The ISIS-Fish model of the Eastern English Channel accounts for technical interactions between 8 species targeted by 17 fleets. It computes fishing mortality based on effort per fleet and métier. We used this model to estimate FMSY and EMSY (the effort needed to reach FMSY) for each species using an effort multiplier instead of an F multiplier assuming fishers keep the same strategy even if stock availabilities change. We then replicated the experience using a fishing behavior model that allows fishers to redirect their effort toward the most profitable métiers as species availabilities change. The results evidenced that the levels of effort needed to maximize species catch ranged between 0.3 and almost 4 times the current level of effort in the fishery depending on the species. The shape of the relationship between effort and fishing mortality were examined and proved to be globally non-linear and diverse. This may explain the high discrepancies obtained between Fmsy computed with ISIS-Fish and ICES estimates available for plaice and sole, ISIS-Fish Fmsy being considerably higher than ICES Fmsy. Differences may also arise from the stock recruitment relationship and the simulation duration (35 years for ISIS-Fish and not until equilibrium). The results obtained with the fishing behavior module indicate higher or identical catch at MSY but the effort required is the same to reach the maximum is the same. EMSY thus appears robust to fishing behavior when FMSY is not.

The perspective of the work is to use the model is order to identify patterns of fishing effort distribution (among métiers or in course of the year) that will bring all species closer to their MSF simultaneously. This can be done using simulation design exploring the effort space or using optimization algorithm providing that an appropriate objective function is found.

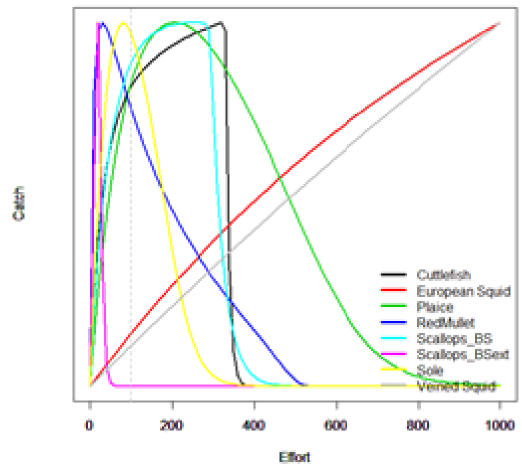


Figure 1. Catch as a function of effort multiplier for the 8 species in the simulations with constant strategies.

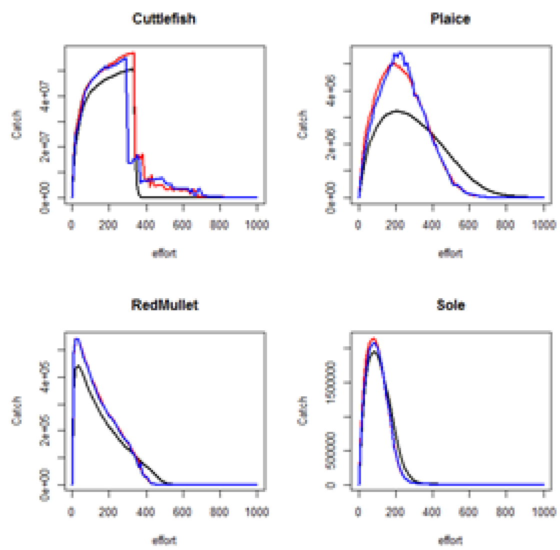


Figure 2. Relationship between catch and effort with constant strategy (black) and using 2 values of opportunism in the behaviour model.