An Integrated Global-to-Regional Scale Workflow for Simulating Climate Change Impacts on Marine Ecosystems

Kelly Ortega-Cisneros[1], Denisse Fierros-Arcos[2], Max Lindmark[3], Camilla Novaglio[2], Phoebe Woodworth-Jefcoats[4], Tyler D. Eddy[5], Marta Coll[6,7], Elizabeth Fulton[8,9], Ricardo Oliveros-Ramos[10], Jonathan Reum[11], Yunne-Jai Shin[10], Cathy Bulman[8], Leonardo Capitani[12,13], Samik Datta[14], Kieran Murphy[2,15], Alice Rogers[16], Lynne Shannon[1], George A. Whitehouse[17], Ezekiel Adekoya[2,18], Beatriz S. Dias[19], Alba Fuster-Alonso[6], Cecilie Hansen[20], Berengere Husson[20], Vidette McGregor[14], Alaia Morell[10,21], Hem-Nalini Morzaria Luna[22,23], Jazel Ouled-Cheikh[6,24], James Ruzicka[4], Jeroesupplemn Steenbeek[7], Ilaria Stollberg[2], Roshni C. Subramaniam[8,9], Vivitskaia Tulloch[25], Andrea Bryndum-Buchholz[5], Cheryl S. Harrison[26], Ryan Heneghan[27], Olivier Maury[10], Jacob Schewe[28], Derek P. Tittensor[18], Howard Townsend[29], Julia Blanchard[2]

[1] Department of Biological Sciences, University of Cape Town, Cape Town, South Africa

[2] Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania, Australia

[3] Swedish University of Agricultural Sciences, Department of Aquatic Resources, Institute of Marine Research, Turistgatan 5, 453 30, Lysekil, Sweden

[4] Pacific Islands Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Honolulu, HI, USA

[5] Centre for Fisheries Ecosystems Research, Fisheries & Marine Institute, Memorial University, St. John's, NL, Canada

[6] Institute of Marine Sciences (ICM-CSIC), Barcelona, Spain.

[7] Ecopath International Research Association (EII), Barcelona, Spain.

[8] CSIRO Environment, Hobart, Tasmania, Australia

[9] Centre for Marine Socioecology, University of Tasmania, Hobart, Tasmania, Australia

[10] MARBEC, Univ Montpellier, IRD, CNRS, Ifremer, Montpellier, France

[11] NOAA Alaska Fisheries Science Center, Seattle, WA, USA

[12] Swiss Federal Institute for Forest, Snow and Landscape Research (WSL, Birmensdorf, Switzerland

[13] Swiss Federal Institute of Aquatic Science and Technology (Eawag), Dübendorf, Switzerland

[14] Population Modelling Group, National Institute of Water and Atmospheric Research, Wellington 6021, New Zealand

[15] The Australian Centre for Excellence in Antarctic Science, University of Tasmania, Hobart, Australia

[16] School of Biological Sciences, Victoria University of Wellington, New Zealand

[17] Cooperative Institute for Climate, Ocean, and Ecosystem Studies, University of Washington, Seattle, WA, USA

[18] Department of Biology, Dalhousie University, Halifax, Nova Scotia, Canada

[19] College of Fisheries and Ocean Sciences, University of Alaska Fairbanks, AK, US

[20] Institute of Marine Research, Ecosystem processes group, Pb 1870 Nordnes, 5817 Bergen, Norway

[21] Puget Sound Institute, University of Washington Tacoma, Tacoma, WA 98421, USA

[22] Long Live The Kings. 1326 5th Ave #450, Seattle, WA 98101

[23] Visiting scientist. Northwest Fisheries Science Center, NOAA-Fisheries. 2725 Montlake Blvd. East. Seattle, WA, 98112

[24] Institut de Recerca de la Biodiversitat (IRBio) and Departament de Biologia Evolutiva, Ecologia i Ciències Ambientals (BEECA), Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain

[25] Basin-scale Events to Coastal Impacts (BECI), North Pacific Marine Science Organization (PICES), Sidney, BC, Canada

[26] NOAA Fisheries Office of Science and Technology, Silver Spring, MD USA

[27] Australian Rivers Institute, School of Environment and Science, Griffith University, Nathan, Queensland, Australia

[28] Potsdam Institute for Climate Impact Research, Member of the Leibniz Association, 14473 Potsdam, Germany

[29] Department of Ocean and Coastal Science and Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana, USA

Corresponding author: Kelly Ortega-Cisneros (kelly.ortegacisneros@uct.ac.za)

Key Points:

- Develops a standardised protocol for detecting past ecosystem changes and simulating climate impacts by regional marine ecosystem models.
- Details tools such as the Regional Climate Forcing Data Explorer Shiny application to access, visualise, and process climate forcing variables.
- The protocol and tools are flexible and can be applied to the different marine ecosystem model types included in FishMIP.

(The above elements should be on a title page)

Abstract (250 max)

As the urgency to evaluate the impacts of climate change on marine ecosystems increases, there is a need to develop robust projections and improve the uptake of ecosystem model outputs in policy and planning. Standardising input and output data is a crucial step in evaluating and communicating results, but can be challenging when using models with diverse structures, assumptions, and outputs that address region-specific issues. We developed an implementation framework and workflow to standardise the climate and fishing forcings used by regional models contributing to the Fisheries and Marine Ecosystem Model Intercomparison Project (FishMIP) and to facilitate comparative analyses across models and a wide range of regions, in line with the FishMIP 3a protocol. We applied our workflow to three case study areas-models: the Baltic Sea Mizer, Hawai'i-based Longline fisheries therMizer, and the southern Benguela ecosystem Atlantis marine ecosystem models. We then selected the most challenging steps of the workflow and illustrated their implementation in different model types and regions. Our workflow is adaptable across a wide range of regional models, from non-spatially explicit to spatially explicit and fully-depth resolved models and models that include one or several fishing fleets. This workflow will facilitate the development of regional marine ecosystem model ensembles and enhance future research on marine ecosystem model development and applications, model evaluation and benchmarking, and global-to-regional model comparisons.

1 Introduction

Climate change is one of the key drivers drastically altering marine and terrestrial ecosystems at rates faster than ever previously recorded (Jaureguiberry et al., 2022; Pörtner et al., 2021). The impacts of climate change differ among regions of the world. Consequently, regionally focused models are needed to meet the needs of considering the effects of climate change at the scales necessary to address the system specific details. Currently, model-based studies project major marine biomass decreases in the tropics by the end of the century, while other areas, such as the Arctic, are expected to experience biomass increases or distribution shifts of economically important species (Cheung et al., 2010; Lotze et al., 2019; Palacios-Abrantes et al., 2022; Rogers et al., 2020; Tittensor et al., 2021). However, the high uncertainty related to these projections can preclude their uptake in decision-making and adaptation planning. Standardised model handling and reporting can help address this by facilitating multi-model comparisons, but also by creating a systematic and repeatable process for those interested in models or their outputs to interact with model products.

Model intercomparisons have been extensively used in climate science to quantify uncertainty in model estimates and projections (Wallach et al., 2016). Their use has been extended to agriculture (Rosenzweig et al., 2013), fisheries and marine ecosystems (Blanchard et al., 2024; Pethybridge et al., 2020; Tittensor et al., 2018), and other sectors (Frieler et al., 2024; IPCC, 2023; Rocklöv et al., 2021). The Fisheries and Marine Ecosystem Model Intercomparison Project (FishMIP) uses ensembles of marine ecosystem models to 'better project the long-term impacts of climate change on fisheries and marine ecosystems and support policy development and long-term planning at the global and regional scales' (Novaglio et al., 2024; Tittensor et al., 2018). As part of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP), FishMIP has developed several protocols (Blanchard et al., 2024; Tittensor et al., 2018) to provide a standardised, structured approach to comparisons of multiple MEMs with the aim of offering more robust projections of changes in biomass and ecosystem structure globally (Bryndum-Buchholz et al., 2019; Lotze et al., 2019; Tittensor et al., 2021). FishMIP considers both global and regional marine ecosystem models (MEMs), which have been calibrated to observations and are used to make medium- to long-term projections of ecosystem dynamics, structure and functioning under different emissions scenarios (Tittensor et al., 2018). A diverse set of regional modelling frameworks, including Atlantis, Ecopath with Ecosim, Mizer and OSMOSE, participate in FishMIP (Audzijonyte et al., 2019; Christensen et al., 2014; Christensen & Walters, 2004; Shin & Cury, 2001). However, due to the patchy global coverage of FishMIP regional MEMs and ensembles, regional extractions of global MEM outputs have often been used to inform on potential biomass change in data-limited areas (Cinner et al., 2022; Tittensor et al., 2018). While such extractions can fill in the knowledge gap, there remains uncertainty as to appropriate ranges of application in terms of system specific characteristics and spatial scale (Eddy et al., this issue).

To date, the focus of FishMIP has mostly been on global MEMs due to their similar spatial coverage, scientific purposes -they have been developed to address climate impact issues by linking to Earth System Models (ESMs), and focus on very similar broad emergent issues in fisheries and ecology. On the other hand, regional models were generally not designed to couple directly to ESMs and tend to be much more specific in terms of objectives, temporal and spatial scales, and have primarily focussed on fisheries issues. This makes regional models much more heterogeneous in content and configuration, and harder to standardise and intercompare. Thus, there is a need to develop a framework tailored to implementing modelling protocols in practice by regional model types within FishMIP. In particular, the standardisation of input and output data is a crucial step in model intercomparisons (Bahlburg et al., 2023; Tittensor et al., 2018) and this is a challenge for models with different structures, assumptions and outputs representing diverse ecosystems and fisheries worldwide. Here we develop an implementation framework and workflow that will guide and improve the implementation of modelling experiments by regional MEMs, thus minimising barriers to entry and thereby increasing the number of regional models performing simulations in a coordinated and standardised manner. FishMIP's vision for regional models includes (i) performing regional-global model comparisons to assess global model reliability and bias for data-limited regional applications, and (ii) fostering regional model ensembles to support case studies. Standardising the climate and fishing effort forcings across regional and global models will facilitate comparisons of MEM outputs, and evaluate the applicability of global models to predict future outcomes in data-poor regions (see Eddy et al., this issue).

This paper aims to present an overview of the approaches used by the different types of FishMIP regional MEMs in conducting climate-impact simulations, and to describe an implementation framework to foster future intercomparisons of MEMs and to ensure they produce assessments that can support policy. The ISIMIP 3a (Frieler et al., 2024) and FishMIP 3a (Blanchard et al., 2024) protocols are used here as a basis for testing the applicability of developing an implementation framework for regional MEMs in FishMIP. FishMIP 3a is the first of the two tracks of the current FishMIP simulation framework (FishMIP 2.0), which addresses the lack of standardised historical fishing data and future fisheries scenarios, and evaluates models against observations before carrying out future projections (Blanchard et al., 2024). "Track A" (FishMIP 3a) focuses on the detection of past climate and fishing impact on historical biomass and catch trends (Blanchard et al., 2024). The goal of this study is to translate the FishMIP 3a protocol into a workflow with practical steps for modelling groups to implement and ultimately facilitate and enable a comparative analysis of MEM outputs within and across a wide range of regions.

2 Materials and Methods

2.1 Marine Ecosystem Model Types in FishMIP

To date, FishMIP includes four regional marine ecosystem modelling frameworks: Atlantis, Ecopath with Ecosim (EwE), Mizer/therMizer and OSMOSE. In addition, EcoTran (Ruzicka et al., 2016) and Models of Intermediate Complexity for Ecosystem Assessments (Plagányi et al., 2014; Tulloch et al., 2019) have recently joined FishMIP. These modelling frameworks are vastly different in model type, representation of species and ecosystem processes, and inclusion and parameterisation of physiological processes affected by climate variables and fishing, among others (Table 1, Tittensor et al., 2018). There is also great heterogeneity in terms of the input data requirements of each model (e.g spatial and vertical resolution). Common key forcings used by regional MEMs are sea water temperature and primary production/plankton biomass (Table 1), and thus these are considered the standard environmental input forcings used by regional MEMs. However, MEMs use a variety of other environmental data as forcing and can include alternative forcings such as oxygen and pH, and even sea ice. Within FishMIP, several EwE models have only used Net Primary Production as climate forcing in the past and have bias-corrected the ESM forcings using the delta method described in Eddy et al. (this issue). A description of the forcings used by each regional modelling framework participating in FishMIP can be found in Table 1 (also see Tittensor et al., 2018; Eddy et al. this issue).

Previous rounds of FishMIP simulations were conducted using outputs from the Coupled Model Intercomparison Project (CMIP) 5 and 6 (O'Neill et al., 2016; Taylor et al., 2012). Details can be found in Tittensor et al. (2018) and Blanchard et al. (2024). FishMIP models, consistent with most MEMs, evaluate the effects of a changing environment on species and ecological processes and use this information to estimate the ecosystem impacts of climate change, while several also include fishing impacts. A major source of uncertainty when projecting climate impacts on marine ecosystems comes from differences in assumptions and structures about the implementation of temperature effects among MEMs (Heneghan et al., 2021; Reum et al., 2024). Some differences between the MEMs in FishMIP include the number of species, functional groups, or size classes affected by temperature changes and the processes affected by temperature and primary production (Table 1). Because of this diversity and the growing number of regional MEMs joining FishMIP (Figure 1), here, we describe an implementation framework and workflow as to how regional MEMs can implement the FishMIP 3a protocol and provide examples of three case studies.

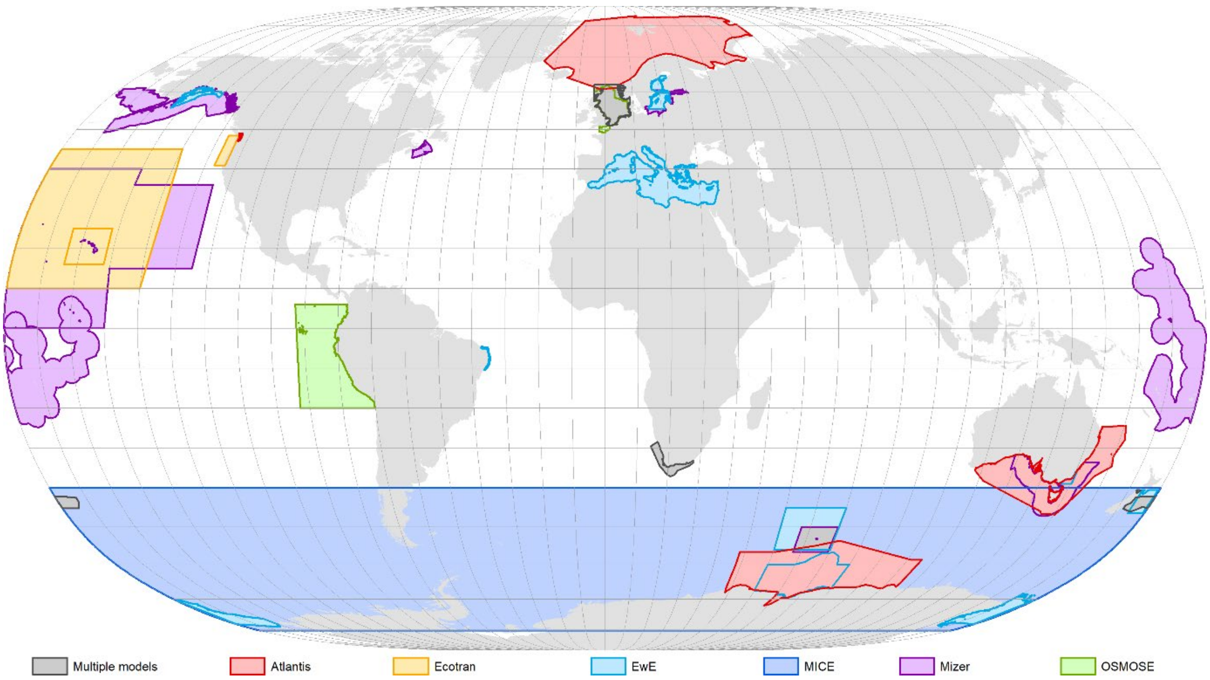

Figure 1. Regional marine ecosystem models participating in FishMIP.

2.2 Simulation workflow

The proposed workflow allows modellers to identify and process the climate model variables of interest, calibrate models to observed data, conduct simulations and contribute outputs to FishMIP and ISIMIP under the standardised FishMIP protocols. The workflow aims to lower the barriers to entry to FishMIP and enable more models to join and perform standardised simulations. The workflow was developed by the FishMIP regional modelling team following best practices for multi-model comparison (e.g. den Boon et al., 2019), and incorporates the experience and knowledge of experts covering all the regional model types included in FishMIP.

Here, the protocol 3a of ISIMIP (Frieler et al., 2024) and FishMIP 2.0 (Blanchard et al., 2024) is used as the basis to provide an implementation framework for regional MEMs. Protocol 3a is aimed at attribution of past changes in marine ecosystems and model evaluation (Blanchard et al., 2024). The latest advancements and efforts conducted by FishMIP to further expand the geographical representation of regional models in FishMIP are also showcased.

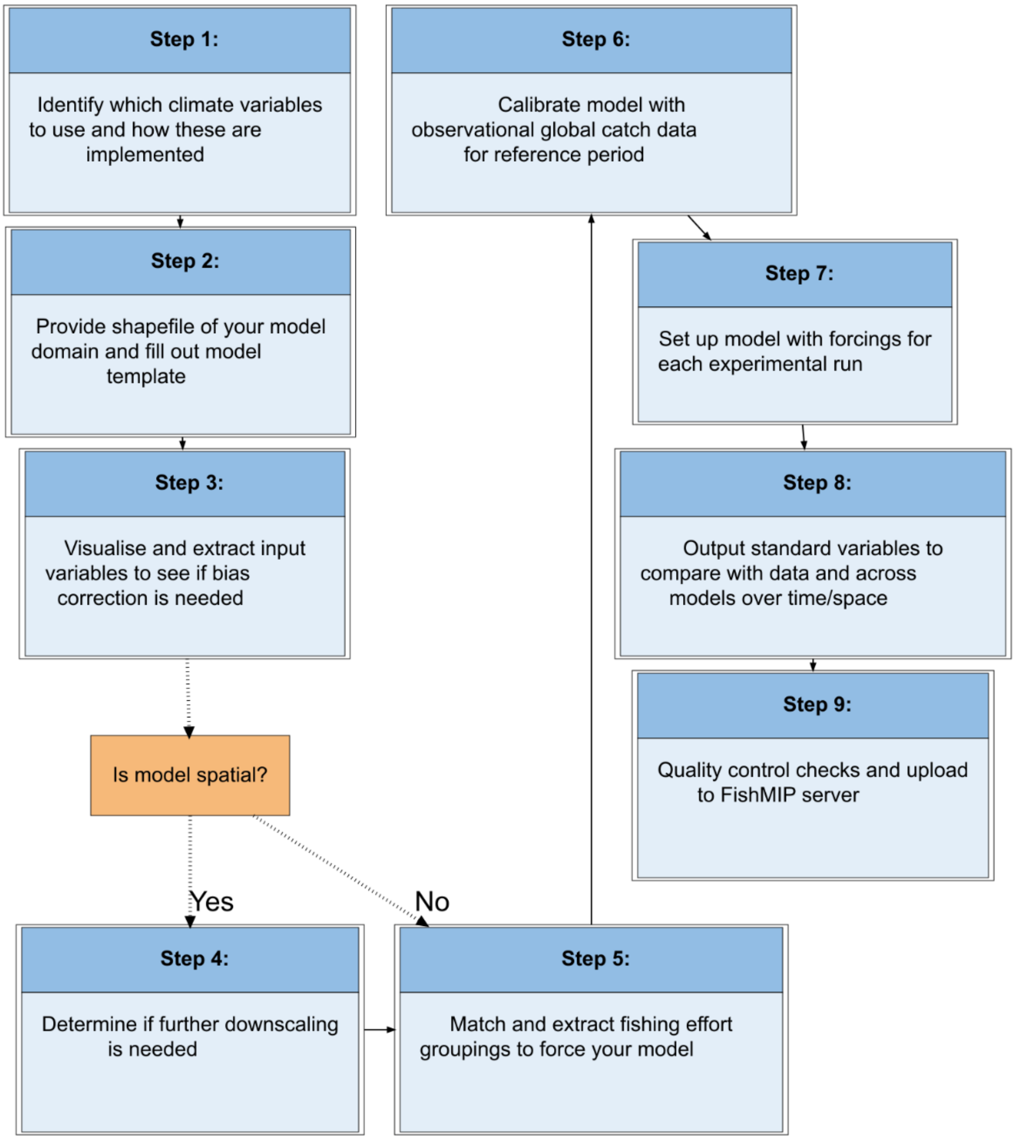

Figure 2. Regional simulation workflow that integrates standardised global forcings with required regional marine ecosystem model inputs. Steps are described in detail below.

Step 1: Identify which climate model variables to use and how these are implemented

Climate forcings are available from ISIMIP, hosted at the German Climate Computation Center (DKRZ) server and the ISIMIP data repository in NetCDF format. ISIMIP has developed tutorials and an Application Programme Interface to access the climate forcings from the DKRZ server. FishMIP has also developed a tutorial on accessing the climate forcings from ISIMIP.

For the FishMIP 3a protocol, oceanic forcing data is derived from the coupled physical and biogeochemical ocean models developed by the Geophysical Fluid Dynamics Laboratory (GFDL): Modular Ocean Model version 6 (MOM6) and Carbon, Ocean Biogeochemistry and Lower Trophics version 2 (COBALTv2). The GFDL-MOM6-COBALT2 model (hereafter GFDL hindcast) was forced by the Japanese 55-year atmospheric reanalysis JRA-55 (Tsujino et al., 2018) and it includes dynamic, time-varying river freshwater and nitrogen inputs that simulate the observed increase in nitrogen loading over the historical period, which is especially important for coastal marine productivity and not regularly included in ESMs (Liu et al., 2021). The FishMIP 3a protocol also makes use of a parallel GFDL-MOM6-COBALT2 simulation without increasing nutrient loading, to test the sensitivity of the FishMIP models to this forcing (hereafter the control). GFDL-MOM6-COBALT2 outputs were regridded to a regular 0.25° and 1° horizontal resolution grid, while preserving vertical resolution. All regional MEMs use forcings at 0.25° horizontal resolution.

A complete list of oceanic climate-related variables available from GFDL-MOM6-COBALTv2 can be found in Frieler et al., 2024 (Table 8) and on the FishMIP 3a protocol. As mentioned above, regional MEMs commonly use sea temperature, primary productivity and plankton biomass to force their models, but differ in the representation of sea temperature and primary production effects (Table 1). For instance, sea temperature can affect different processes in the different regional MEMs, such as movement of ecological constituents in some models (e.g., Atlantis, Ecospace and OSMOSE), while mortality and/or assimilation efficiency can be affected by temperature in EwE, Bioen-OSMOSE and Atlantis. Regarding primary production and plankton biomass, most MEMs can use plankton biomass derived from ESMs and override the plankton dynamics within the MEM. Table 1 summarises how temperature and primary production/plankton biomass forcings are implemented in the FishMIP regional MEMs.

Step 2: Provide shapefile of your model domain and complete model template

As per Step 1, modellers have the option to (i) access climate forcings directly from the DKRZ server or the ISIMIP repository or (ii) provide model spatial boundaries (shapefile or bounding box) for the regional modelling team to extract all climate variables available in GFDL-MOM6-COBALTv2 (Table 8 of Frieler et al., 2024, FishMIP GitHub page). The creation of Python scripts to complete this step has streamlined the process into a standardised format for the 34 participating FishMIP regional models (Fig. 1, as of April 2024). The Python scripts developed for regional data extraction are publicly available in the FishMIP GitHub repositories. Regional climate forcings are also publicly available at the University of Tasmania THREDDS server.

Modellers were required to document how the climate and fishing forcing were integrated into their models to ease the quantification of uncertainties due to differences in model structure and assumptions and the analysis of ensemble MEM projections (den Boon et al., 2019). This includes the resolution of the climate forcing used, the environmental forcings equations used, and which ecological process each forcing affects, the fishing forcing set-up–e.g., fishing mortality rates, selectivity and catchability estimates, and how fishing gears and functional groups targeted were aggregated–as well as details on model calibration (Supplementary Information II). Because models involved in FishMIP evolve through time, questionnaires with information about regional marine ecosystem models are stored in the FishMIP GitHub. Information on the model templates also feeds the model documentation on the ISIMIP website.

Step 3: Visualise and extract input variables to see if bias correction is needed

Visual comparison of climate forcings from ESMs against observations for the region of interest is necessary to determine whether bias correction is required. To improve the accessibility of climate data to different regional modelling teams and ease the processing of ocean forcings, FishMIP is currently focused on (i) improving the workflow before FishMIP protocols are finalised and modelling experiments are run and (ii) developing tools that contribute to these modelling efforts (Novaglio et al., 2024). The development of the 'Regional Climate Forcing Data Explorer' Shiny app (Fig. 3, left panel) represents one of these steps. The shiny app shows climatological means from 1961–2010 (historical period of the 3a protocol) as maps, and spatial averages as time series, for 37 ocean variables available in GFDL-MOM6-COBALTv2 for the regional models currently participating in FishMIP. These ocean forcings can be downloaded for each model region for use as inputs by regional MEMs.

Climate model outputs are known to have systematic biases, which can preclude their direct use for regional climate-impact and vulnerability assessments (Casanueva et al., 2020). A number of bias correction methods have thus been developed to correct the climate model outputs using observations at regional scales (Casanueva et al., 2020 and references therein). These methods differ in complexity and can be trend-preserving or not, correct the mean to univariate or multivariate metrics, and robustly adjust extreme values (Casanueva et al., 2020; Lange, 2019). The implications of bias correction include possible impacts on magnitudes, signals or trends (Oliveros-Ramos et al. in revision). For FishMIP 3a, regional modellers observed differences in sea temperature and primary production between the GFDL hindcast (1961-2010) and those derived from regional ocean models or observations (Fig. 3, see section 4.1 for an example of three case study areas). These temperature differences resulted in having species outside their thermal tolerance ranges causing some of them to collapse during pilot historical model runs. It was therefore decided to perform bias correction on the GFDL outputs. The delta method for calibrating the mean (see Supplementary Information I) to observations was chosen due to its relative simplicity and applicability (Marshall et al., 2017; Pozo Buil et al., 2023).

The selection of the dataset used to perform bias correction is of utmost importance as previous studies found that bias correction methods strongly rely on the reference dataset used for calibration. We used the Word Ocean Atlas 18 (WOA) (Garcia et al., 2019; Locarnini et al., 2018) because this is a comprehensive, quality controlled dataset based on ocean profiles data from 1955 to 2017, providing gridded climatology fields for temperature, salinity, oxygen, among other variables. The WOA datasets have been extensively used for bias correction purposes (e.g., Séférian et al., 2013 (WOA09); Fu et al., 2022 (WOA18)). Global reanalysis products such as GLORYS were not used at this stage because their temporal range does not match the time span of the ISIMIP and FishMIP protocol 3a (i.e. GLORYS starts in 1993, and the FishMIP protocol starts in 1961). A list of sequential steps to perform bias correction on sea water temperature can be found in Supplementary Information I. Those steps can also be used for variables such as salinity and oxygen.

Different approaches have been used to bias correct plankton biomass and primary productivity within FishMIP regional MEMs (Table 1). A common approach involves using the delta method to adjust ESM outputs and force primary production (Eddy et al., this issue) and the growth of plankton groups (Rovellini et al., 2024).

Step 4: If spatial: determine if further downscaling is needed

Given the complexity of downscaling approaches and the need to evaluate their performance on a regional basis, we have not yet standardised the statistical downscaling approach to be used in this implementation framework (other than performing bias correction). ISIMIP has a bias correction and statistical downscaling protocol, which has been applied to atmospheric climate data and it is likely not directly transferable to oceanic variables (Lange, 2019). If this step needs to be carried out by regional modellers, we advise the modeller to choose a statistical downscaling approach that performs best for their region and use the WOA18 dataset and the time periods specified in step 3 (see Supplementary Information I) to perform the downscaling and to ensure consistency with this implementation framework. We acknowledge that standardising the choice of a statistical downscaling method is an area that warrants further attention within FishMIP.

Major differences have been found between low-resolution ESM outputs and highly resolved downscaled projections at a regional scale (Melsom et al., 2009; Skogen et al., 2018). When forcing the Nordic and Barents Atlantis model with an ESM (1° resolution) and a regional ocean model (dynamically downscaled projections at 10 km resolution), a general agreement in future biomass trends and distribution patterns for some species at higher trophic levels were found, but this was not the case for lower trophic level groups (e.g., plankton, mesopelagic and prawns), and for some higher trophic level species such as Northeast Arctic cod (*Gadus morhua*). These differences indicate that highly resolved forcings are needed in studies focused on coastal systems (as is the case for most regional

MEMs) and/or representing finer-resolution processes. However, downscaled climate forcings, especially dynamically downscaled, are not available for most regions of the world, nor the full set of climate scenarios, and this represents a challenge for regional climate-impact assessments (Kristiansen et al., 2024; Pozo Buil et al., 2021).

OSMOSE-Humboldt is the only FishMIP regional model type that has performed statistical downscaling using methods other than the delta method commonly used to perform bias correction (Step 3). Oliveros-Ramos et al. (2023) evaluated 19 nested statistical downscaling models describing the relationship between empirical distributions of historical modelled and observed SST using ten indicators of predictive performance for model selection. They did not find a single statistical downscaling model that performed better than all others across regions. Instead, model performance varied across regions, indicating that these approaches should be evaluated on a case-by-case basis. The 'Gridded time series analysis' R package implements the statistical downscaling models described in Oliveros-Ramos et al. (2023). Statistical downscaling does not require the use of high-performance computing (as required by dynamical downscaling), and this is extremely important as lower requirements for technical skills and computational capacity may result in a higher adoption rate within the modelling community, especially among researchers starting in this field. This is one approach currently being evaluated for future use within FishMIP.

Step 5: Match and extract fishing effort groupings to force your model

For FishMIP protocol 3a, global fishing effort time series were made available to FishMIP modellers (Blanchard et al., this issue), and future scenarios are being developed for Phase 3b (Maury et al., this issue). This represented a significant step forward, as this allowed global modellers to represent historical fishing impacts, which many global MEMs were not able to include before such global data were available. Regional models did include fishing, and in most cases, used statistics from government agencies or regional advisory organisations. Regional modellers generally consider this regional fishing effort information more accurate, and several discussions were held to find the best way to use global effort data developed for protocol 3a to standardise fishing forcing between global and regional MEMs and improve the comparability of their outputs.

The fishing effort data provided by FishMIP (hereafter called global effort data) was derived from Rousseau et al. (2024) and consists of 16 gears or fleets and a total of 29 functional groups (Table 2). Fishing effort data used to force regional models and fishery catch data (Watson & Tidd 2018) used for model calibration were processed and extracted for each regional MEM by the FishMIP coordination team and are publicly available in the FishMIP THREDDS server. More details on the regional extraction of catch and effort data can be found in Blanchard et al. (2024).

Most regional models include at least some of their ecological components at the species level, or at least at taxonomic resolutions finer than reported in aggregated global statistics. Consequently, it was necessary to make some assumptions on how to split the global effort and catch data by fleet and functional group to match the taxonomic resolution of the regional model considered. Regional effort and catch time series (where available) are to be used in combination with the global data to inform the processing assumptions (e.g. disaggregation of effort by functional groups into species). Careful consideration and a preliminary analysis of the FishMIP effort data for some model regions highlighted important inconsistencies with effort data from regional management authorities and other local sources commonly used by regional modellers (see section 4.2). Inconsistencies were mostly due to the nature of the global data, which is global in coverage but less detailed and reliable at the regional scale. To address this issue, three sensitivity tests are proposed for the implementation of the global data:

1. Global effort data only: If there is a good agreement between the historical trends and magnitude of the global and regional effort data. Modellers implement the global effort data into their regional MEMs following the procedure described in Supplementary Information 1.
2. Bias-correction of the global effort using regional data: If there are differences between the historical trends and magnitudes of the global and regional effort data for some fleets. Modellers

can use the global effort data for those fleets showing reasonable historical trends and use their regional effort/mortality to correct the global effort forcing for those that do not.

3. Regional effort data only: If there is little agreement between the historical trends and magnitude of the global and regional effort data. Modellers should use their regional effort/mortality to perform simulations as per their baseline models. Modellers are requested to describe the differences between these datasets to justify the use of regional data and to ensure improvements are made in future. This will also allow us to evaluate the influence of global vs regional effort forcings on historical model outputs.

Modellers are requested to submit their fishing effort/mortality time series with their simulations. We acknowledge that regional effort and catch time series are often not publicly available as they belong to national government agencies. In those cases, we ask modellers to submit their forcings as relative values if this does not contravene the access conditions under which the data was granted. The sequential steps involved in processing the global effort and catch data to obtain a time series of fishing effort and total catch split by fleet and functional groups can be found in the Supplementary Information I. Code has been provided for worked examples that illustrate this step.

Step 6: Calibrate MEM outputs with observational global catch data for reference period

Calibrating MEM outputs to observational data is a computationally- and time-intensive process. For some models (EwE, Mizer), it may be feasible to recalibrate models with all climate and fishing forcings since specific protocols exist. We have provided catch data extracted for each regional shapefile to facilitate this step in cases where no other data are available (step 5) or where experimental design necessitates. Even though the 3a experiments extend to 2010, the catch time series extends up to and including 2004. Later years (2005-2010) must not be used in calibration because we have retained the last six years of the catch data for predictive skill assessment across models.

In cases where recalibration cannot be carried out, we still encourage modellers to submit their runs and compare them to the outputs of their baseline calibrated runs, including inputs. In this case, we ask modellers to submit the results of their baseline model runs. It may, in some cases, be appropriate to carry out a statistical post-hoc adjustment of simulations based on the discrepancy of the two runs. Another possibility is simply to provide the non-calibrated runs with a clear indication in the model template that recalibration was not carried out. In these cases, an analysis of relative changes may still be performed, keeping in mind that the non-calibrated model may have limited performance when capturing observed historical changes for the system in question.

In all cases, we expect modellers to carry out "sanity checks" of their models. This is step 0 of the Hipsey et al. (2020) framework. This involves ensuring that processes and rates in each MEM are plausible and sensible. We then suggest using a subset of the model skill metrics to assess how well the MEMs forced with the global effort data compared to the original MEM calibrated with regional effort/mortality data. A minimum set of suggested metrics and plots include bias and correlation of time series of catches and, if observations are available, biomasses for key functional groups and species in the model. We ask modellers to submit the data and all data sources (when those are publicly available) used in this step if different to what has been provided and detailed in steps 3 and 5. When this is not feasible, possibly due to permissions, relative time series and summary statistics should be provided.

A toolbox is being developed to analyse and compare spatial model outputs within an integrated and standardised workflow and calculate a number of skill metrics (i.e. MapCompR). MapCompR provide functions to i) compare spatial maps from different species, ii) compare spatial maps of the same species obtained with different methods, and iii) analyse model predictions.

Step 7: Set up MEMs with forcings for each experimental run

The FishMIP protocol 3a consists of four model experiments and eight scenarios, with different combinations of climate and human forcings (see Table 1 of the FishMIP 3a protocol). A model experiment is a set of model simulations with a particular goal (e.g. model evaluation), while a scenario is a particular setting for climate and human forcing drivers (e.g. fishing). The two core experimental runs aim to evaluate the impacts of climate with time-varying river input forcing at 0.25° resolution (step 3), with and without fishing (step 5). Two optional but preferred runs were set up to estimate the sensitivity of model outputs to riverine influx (ctrlclim, input forcings held at 1955 values throughout the simulations). This model experiment is also run with and without fishing.

Two additional experiments were also set up in the FishMIP 3a protocol, aiming to understand the impacts of resolution on model outputs, and use climate forcings at a 1° resolution with exactly the same set-up listed above for the core and preferred runs. In translating the FishMIP 3a protocol to a regional context, we decided to focus on the experiments using 0.25° resolution forcings (i.e. the core runs) due to the finer resolution needed to force regional models.

Step 8: Output standard variables to compare with data and across models over time/space

The FishMIP protocol 3a lists all the mandatory and optional model outputs to be provided by modellers (Table 9, FishMIP protocol 3a), including the variable specifiers. We request that modellers report what species and species groups were allocated to the different output variables (Table 9, FishMIP protocol 3a) in the model templates (step 2). Regional modellers should submit their spatial outputs as NetCDF files, while outputs from non-spatial regional MEMs can be saved as .csv files.

The optional outputs include indicators such as the biomass and catch of different size classes of pelagic and demersal fish. These outputs are highly relevant at the regional scale as they can be directly linked to system specific species of ecological and economic importance. The mandatory and optional outputs will also allow the estimation of ecosystem indicators (Coll et al., 2016; Shin, Bundy, et al., 2010; Shin, Shannon, et al., 2010), which are regularly calculated in regional modelling studies in a number of regions and allow for a further point of comparison. These indicators include species-based, size-based and trophodynamic indicators that have already been compared across regional MEMs and ecosystems in the frame of the IndiSeas working group (Fu et al., 2019; Ortega-Cisneros, Shannon, et al., 2018; Reed et al., 2016; Shin et al., 2018). Depending on the scenarios and forcings considered, a subset of indicators could be used that are the most sensitive, responsive and specific to changes in drivers. For example, Shin et al. (2018) showed that among the IndiSeas indicators tested, mean fish length had the more specific response to changes in plankton biomass, while total catch/biomass ratio was more specific to changes in fishing pressure. Recent sensitivity and uncertainty analyses can be used to identify the indicators that are more robust to uncertainties (Luján et al., 2024). Along the lines of Luján et al. (submitted), a standardised protocol could be developed in the future for the FishMIP MEMs to identify a common set of indicators that are robust to uncertainties in model parameterisation.

Step 9: Quality control checks and upload MEM outputs to FishMIP server

There are strict specifications on how to prepare and name MEM outputs for submission to FishMIP. File names consist of a series of identifiers including the regional MEM type, climate forcing, the climate, socioeconomic and sensitivity scenario identifiers, and the variable identifier, region and timesteps. Specific guidelines and instructions can be found on the ISIMIP website and the FishMIP protocol 3a repository.

This is a seemingly trivial but extremely important step to ensure ensemble consistency and expedite analysis. It is crucial that modellers follow closely the formatting guidelines for reporting model outputs to facilitate their analysis within the ISIMIP framework. Regional modellers should use the quality control tool developed by ISIMIP, which allows modellers to check their outputs against the definitions and conventions of ISIMIP protocol before submission. Regional modellers should contact the FishMIP regional modelling team if they have questions about how to format their MEM outputs. Once model outputs are ready for submission, modellers must save them on the upload area (a folder is available for each model region and type) of the DKRZ server.

Applying the framework

The workflow described here (Fig. 2) has been applied to three case study areas-models: the Baltic Sea Mizer, the Hawai'i-based Longline therMizer and the southern Benguela ecosystem Atlantis regional models. Details on these models (e.g. functional groups, fleets, calibration and skill assessment) can be found in Supplementary Information I and the FishMIP GitHub repository. The results below represent a subset of the steps described in the workflow and were selected to illustrate the implementation of the most challenging steps of the workflow and how they can be applied to different MEM types and model regions to illustrate the applicability and flexibility of the workflow.

4 Results

4.1 Case study 1: Climate forcing intermodel comparison

In step 3 of our workflow, our shiny app is available for modellers to extract climate forcings for their region, visualise them and download the variables they need to compare them to standardised global or regional observation datasets. Any region can be selected to visualise and download the 0.25° resolution forcings (see Figure 3A, C, E for sea temperature), then each model may aggregate this data as required. To further assess whether bias correction is required for physical ocean variables (i.e. temperature), a comparison with WOA observations was carried out.

Baltic Sea Mizer model

The Baltic Sea Mizer model uses sea surface temperature as model input, averaged over the whole model domain (Lindmark et al., 2022). A time series of monthly sea surface temperature was acquired from the GFDL hindcast, spanning from January 1961 to December 2010 (Fig. 4B). This hindcast represents the 'climate with observed atmospheric forcing and river input forcing'. Similarly, an average sea surface temperature value was calculated for the control and the WOA datasets. The bias corrected time series (Fig 4B) was compared to the GFDL hindcast to determine if bias correction was needed for this model. The absolute difference between these time series was 0.56 °C, and suggested that bias correction may be needed for this model. Based on the temperature difference between datasets, it is expected that some modelled species may show unexpected behaviour during model simulations.

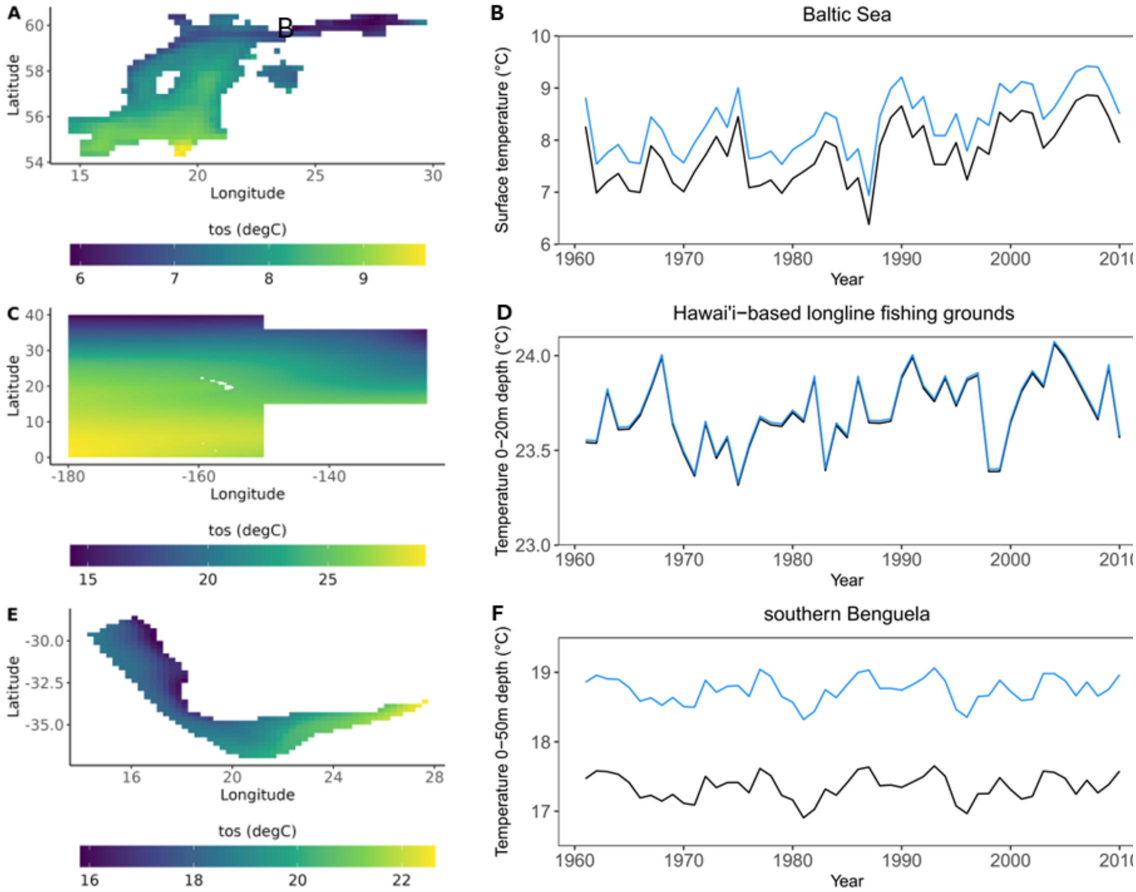

Figure 3. Maps of surface temperature climatological means (1961-2010) calculated from GFDL-MOM6-COBALT2 hindcast for the Baltic Sea Mizer (A), Hawai‘i-based longline fishing grounds (C), and southern Benguela (E) model domains. Time series of bias-corrected (black lines) and GFDL-MOM6-COBALT2 hindcast (light blue lines) sea temperature for the surface within the Baltic Sea Mizer model (B), for the top 20 m within the Hawai‘i-based longline fishing grounds (D) and for the top 50 m of the southern Benguela (F). The different depth intervals used to integrate sea temperature in panels B, D and F reflect the different input forcings used by each model (see section 4.1 for more information). The bias-corrected time series were calculated using the procedure detailed in the Supplementary Information I.

Hawai‘i-based longline therMizer model

The Hawai‘i-based longline therMizer model uses temperature averaged over 18 depth ranges as model input. This model captures species' vertical behaviour and exposure to different depths, and includes temperature at depth ranges from 0–20 m up to 400–1200 m depth (see Supplementary Information I for an explanation of the approach). Eighteen temperature time series (January 1961 to December 2010) were acquired for this model from the GFDL hindcast. Each time series corresponds to the 18 preferred depth ranges for the model species (see Fig. 3D for an illustration of average temperature at 0-20 m depth), while 18 average sea temperature values were calculated for the control and the WOA datasets. The comparison between the GFDL hindcast and the bias corrected time series indicates small absolute differences in temperature (0.013 °C) for the 0–20 m depth range. While the bias was negligible for the 0-20m depth layer for this model, the bias was higher for deeper depths, and simulations (results not shown here) using the GFDL hindcast without bias correction resulted in some species going extinct during the simulations because the GFDL hindcast temperatures fell outside observed temperatures. This highlights the importance of the bias-correction step for some models, specifically those including functional groups with narrow thermal preferences.

Southern Benguela ecosystem Atlantis model

The southern Benguela ecosystem Atlantis model is a spatially explicit model, for which the model area is divided into 18 polygons (Ortega-Cisneros et al., 2017). The model extends to a maximum depth of 500 m, with two depth layers near the coast and four offshore (Fig. 4) and an assumption of an open boundary layer underlying the offshore boxes (1000 m depth). The procedure detailed in step 3 (Supplementary Information I) was followed as was the case for the Baltic Sea Mizer and Hawai'i-based longline therMizer models. For the southern Benguela Atlantis model, this procedure resulted in 59 time series of sea water temperature (1961-2010) from the GFDL hindcast and 59 average sea water temperature data points each for the control and WOA datasets. This was because it was necessary to aggregate the gridded inputs into the 18 spatial polygons used as the spatial configuration for this regional model (instead of one for the whole model area), and then to calculate average temperature for the different depth layers used in this model (Fig. 4). For illustrative purposes, the bias corrected and GFDL hindcast temperature time series for two model polygons of the southern Benguela ecosystem Atlantis model are shown in Figure 4. The difference between these datasets is 1.43°C at the 0-50 m depth layer (Fig. 3F), and increased with depth to 3.48 °C for the 300-500m depth layer. It is therefore expected that using the GFDL-MOM6-COBALTv2 hindcast without bias correction would likely result in several modelled species going extinct during model simulations.

For other spatially explicit models (case-dependent, step 4), comparing them with gridded observed climatologies can help indicate whether further statistical downscaling may also be needed (e.g. Oliveros-Ramos et al., 2023). For this, we recommend following the guidelines provided in step 4.

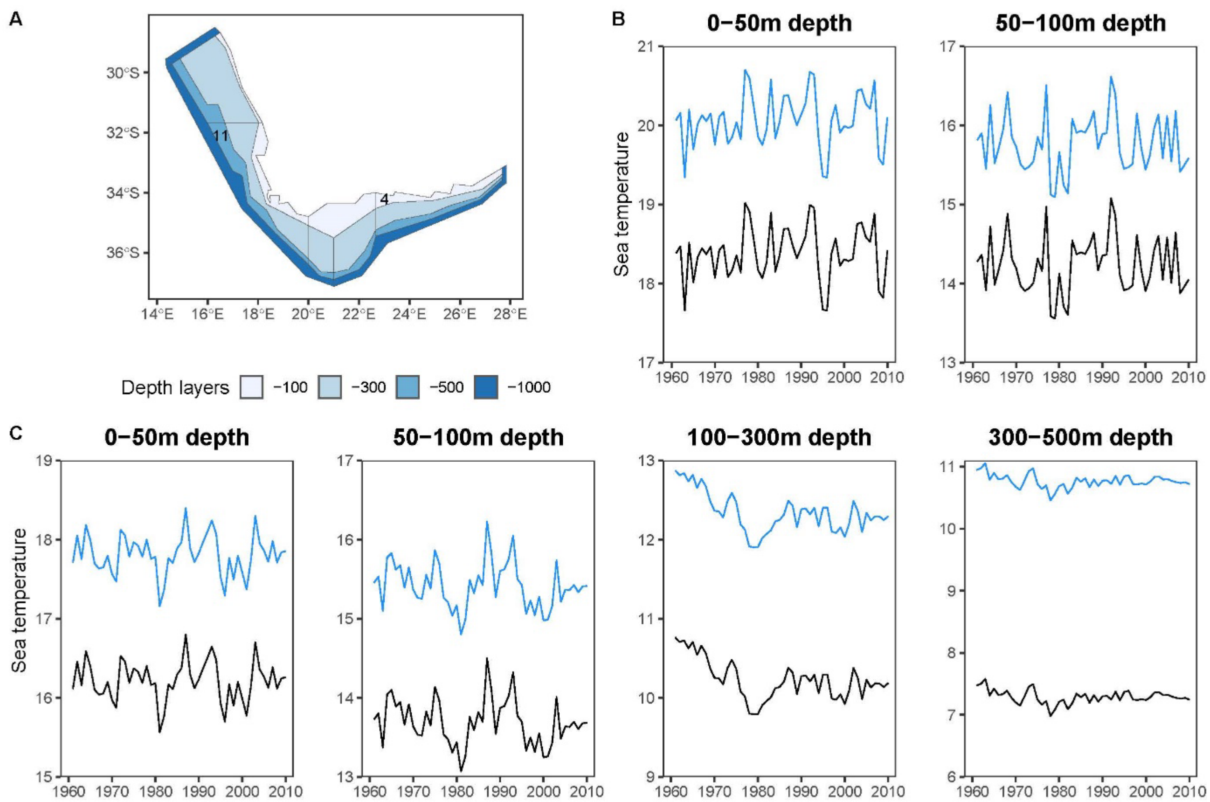

Figure 4. Model geometry of the southern Benguela Atlantis model showing model polygons and depth layers (A). Time series of bias corrected (black) and GFDL-MOM6-COBALTv2 hindcast (light blue) temperatures at different depth ranges for model polygons 4 (18 boxes × 2 depth layers) (B) and 11 (18 boxes × 4 depth layers) (C).

4.2 Case study 2: Fishing effort forcing intermodel comparison

All regional MEMs in FishMIP include fishing impacts. However, they vary in their representation of those impacts, such as the use of fishing effort or mortality, the number of fleets, and the number of functional groups impacted by fishing. Here, we provide an overview of how the global fishing effort was used for our three regional MEMs, including one to several fleets.

All fisheries models are based on the premise that fishing mortality is the product of selectivity × catchability × effort. Only effort was varied in the construction of the fishing forcing, with selectivity and catchability unchanged from the way in which the respective models typically deal with these parameters. In the Baltic Sea Mizer and Hawai'i therMizer selectivity and catchability were set to 1 throughout for both. For the southern Benguela ecosystem Atlantis, catchability is set to 1, and constant age selectivity is used with fishing mortality. For anchovy, age selectivity applies to fish older than six months and for sardine older than one year.

Baltic Sea Mizer model

The Baltic Sea Mizer model required an alternative approach to how fishing was incorporated. This Mizer model consists of three fish species: Atlantic cod (*Gadus morhua*), Atlantic herring (*Clupea harengus*) and European sprat (*Sprattus sprattus*). The original model (Lindmark et al., 2022) was calibrated to stock-level fishing mortalities and did not explicitly include different fleets. The majority of landings of cod stem from the bottom trawl fleet ("Trawl_Bottom"), and the majority of sprat and herring by pelagic trawl fleet ("Trawl_Midwater_or_Unsp") (verified using logbook data and assessment reports from the regional advisory organisation ICES). Therefore, these gears were selected in the initial processing of the global effort data. The effort ("NomActive") was next summed by year and functional group, where cod belongs to "demersal30-90cm" and sprat and herring belong to "pelagic<30cm". A time series of relative global fishing effort was made by dividing the effort by the maximum in the time window 1992–2004. This deviation from the workflow (scaling to maximum rather than mean) was made because the bottom trawl effort was characterised by a few large spikes in effort (two years with fishing efforts larger than 5 standard deviations above the mean). To go from relative fishing effort to fishing mortality in the Baltic Mizer model, the mean difference between the fishing mortality derived from stock assessments and that of the relative effort time series over the time period 1961–2010 was added to the relative time series to correct the global effort forcing. The time series of assessment-derived fishing mortalities and global fishing effort are shown in Fig. 5A-C. The validation compared these time series for cod, herring and sprat through a correlation; the Pearson's correlation coefficient r was -0.203 ($p = 0.156$), 0.497 ($p < 0.0001$) and 0.6 ($p < 0.0001$) for cod, herring and sprat respectively. The model predicted average spawning stock biomass (SSB) (forced with global climate and fishing data) was compared to the average SSB from the assessment in the calibration time window (1992–2004), as in the original publication (Lindmark et al., 2022). The model returns a comparable SSB as the original model for cod and herring (77 vs 56 tonnes, and 600 vs 532 tonnes for the original model and the one forced with global data, respectively), while sprat SSB is nearly half in the simulation with global forcings due to the considerably higher effort in the global effort data. This is partly explained by sprat having higher fishing mortality in the global data (mortalities are on average +0.25 higher than the assessment fishing mortalities) in the calibration time window.

Hawai‘i-based longline (ther)Mizer model

The Hawai‘i-based longline model (Woodworth-Jefcoats et al., 2019) includes the longline fleet ("Lines_Longlines"), hence this fleet was selected in the initial processing of the global effort data. The modelled Hawai‘i-based longline fleet catches 12 model species included in three pelagic ("pelagic<30cm", "pelagic30-90cm", "pelagic>=90cm") and two shark ("shark<90cm", "shark>=90cm") functional groups. The effort ("NomActive") across these five functional groups was aggregated to estimate the total effort of the longline fleet per year, under the assumption that a single longline fleet is catching these functional groups. This assumption is based on the characteristics of the Hawai‘i-based longline fleet.

The catch data used to inform the Hawai'i-based longline model starts in 1995, and thus, a baseline average effort was calculated using the time period 1995–2004. The time series of global effort ("NomActive") for the longline fleet was then divided by the baseline average effort to estimate the relative global fishing effort. The global relative fishing effort was multiplied by 0.2, which is the fishing mortality ($F = 0.2$) used to calibrate the Hawai'i-based longline therMizer model (Woodworth-Jefcoats et al., 2019) to arrive at a time series of fishing mortality values (Fig. 5D). Fishing mortality $F = 0.2$ was used in this model because a fishing mortality close to 0.2 has been estimated for those species with available stock assessments (Woodworth-Jefcoats et al., 2019 and references therein).

The Hawai'i-based longline therMizer model applied the global fishing effort to the functional groups caught by the longline fleet. A validation run was performed using constant fishing mortality ($F = 0.2$) as per the original model (Woodworth-Jefcoats et al., 2019). The validation used a correlation test to compare observed and modelled catch at size for the 12 species targeted in the model. All correlations were significant (max p-value = 0.0028), while the Pearson's correlation coefficient r ranged from 0.296 to 0.922, with a mean of 0.65 and a median of 0.684.

Southern Benguela Atlantis model

The southern Benguela Atlantis model followed the approach detailed in step 5 (Supplementary Information I), as described for the Hawai'i-based longline therMizer model. The southern Benguela Atlantis model (Ortega-Cisneros, Cochrane, et al., 2018; Ortega-Cisneros et al., 2017) includes purse seine, inshore and offshore demersal trawl, mid-water trawl, line and jig fisheries targeting a number of functional groups within the model. The original model was calibrated against biomass and catch time series for key functional groups (Ortega-Cisneros et al., 2017).

The purse-seine fishery, targeting small pelagics, is the largest fishery in terms of landings in South Africa (DFFE, 2023). Therefore, this fleet was selected for the initial processing of the global effort data. First, the effort ("NomActive") for the purse seine fleet ("Seine_Purse_Seine") was filtered. This fleet targets anchovy (*Engraulis encrausicolus*) and sardine (*Sardinops sagax*), and also round herring (*Etrumeus whiteheadi*) in recent years; these species belong to the "pelagic<30cm" functional group in the global effort data. A relative time series of global effort for the purse seine fleet and the "pelagic<30cm" was then estimated using the baseline effort calculated from 1990-2004 (the southern Benguela ecosystem Atlantis model starts in 1990). The conversion from relative fishing effort to fishing mortality was achieved by multiplying the relative effort time series by the annual baseline fishing mortality for anchovy and sardine in this model (Fig. 5E-F). The correlation between the global effort data for the purse seine fleet and the harvest proportion for anchovy and sardine derived from the stock assessment for these species (de Moor, 2021) was estimated as a form of validation. A high and significant correlation was found for sardine ($r = 0.668$, $p < 0.0001$) but not for anchovy ($r = -0.171$, $p = 0.459$).

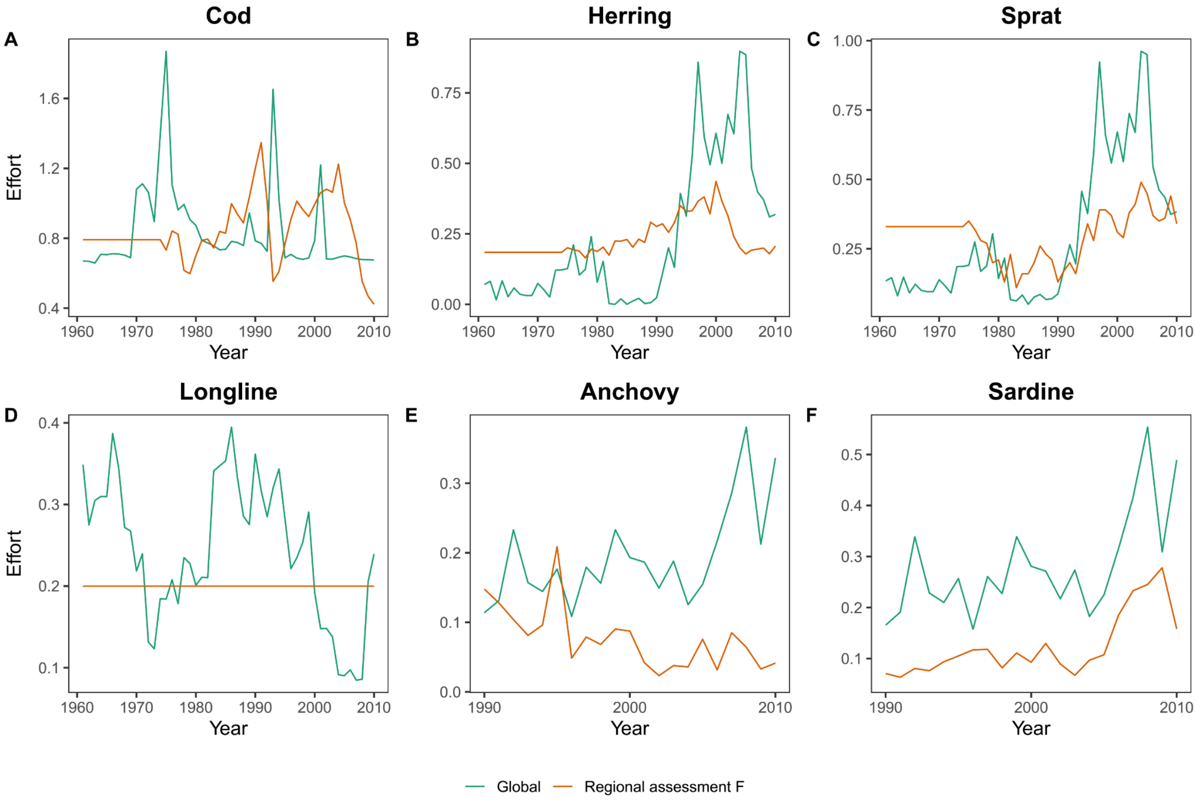

Figure 5. Annual global fishing effort time series for key functional groups compared with regional inputs for the Baltic Sea Mizer (A-C), Hawai'i-based longline therMizer (D) and southern Benguela ecosystem Atlantis (E-F) regional models. Global fishing effort refers to the effort time series calculated using the effort provided by FishMIP and the regional assessment refers to the fishing mortality or harvest proportions derived from stock assessments (see section 4.2)

5 Discussion

Here we described an implementation framework for regional MEMs to participate in comparative analyses as part of FishMIP, across models and a wide range of regions worldwide. Our workflow for setting up regional MEMs for climate hindcasts or projections is flexible enough to apply to a range of MEM types. The case study intercomparison applications of our workflow show that each specific model-region combination has unique requirements that can be accommodated by the extraction tools we have designed. We envisage this workflow will facilitate future research on MEM ensemble development and applications in at least the following ways: 1) regional MEM ensembles, 2) model evaluation and benchmarking (across multiple models/regions), 3) global-regional model intercomparison for regions.

5.1 Regional marine ecosystem model ensembles

The framework presented here provides modellers with a workflow that allows them to process climate and fishing forcings in line with their model requirements and the resources of the modelling team to perform the simulations. Our protocol proved flexible in accommodating MEMs with one fleet (Hawai'i-based longline therMizer model) or several fleets targeting different functional groups (Baltic Sea Mizer and southern Benguela ecosystem Atlantis models). Notably, the availability of the global fishing effort also represents an important step for regions where local fishing effort and mortality are unknown or where records are incomplete, as this will allow regional modellers to represent the impacts of fishing on their MEMs. In addition, the global effort data can be used to represent artisanal fisheries, for which there is limited available data worldwide (Cisneros-Montemayor et al., 2020). It is, however, recommended that the limitations of such an approach (see section 5.4) be clearly communicated to any end-user of such projections (e.g. decision makers) and that global effort data be combined with any available regional information or knowledge from local experts to improve the implementation of the global data into regional MEMs.

We hope the development of this workflow will accelerate and foster comparisons of MEMs across and within regions. For instance, MEM ensembles can be used to conduct experiments and test scenarios in a standardised manner or to perform in-depth evaluations of uncertainty sources in climate projections (e.g. Murphy et al., this issue). The latter is particularly important given the increasing need for MEM outputs to support policy and decision-making, for which regional models should be particularly suited.

5.2 Model benchmarking

Benchmarking is necessary to improve the uptake of MEM outputs and to make them policy-relevant (Frieler et al., 2024). There are several different approaches to benchmarking, ranging from quantifying error to fully conducting uncertainty assessments (Luo et al., 2012; Mackinson et al., 2018; Ogunro et al., 2018). One of the main issues related to improving the reliability and robustness of projections by MEMs is their limited cross-ecosystem validation against historical data (Heneghan et al., 2021; Novaglio et al., 2024), which is true at both global and regional levels. One of the reasons is the limited observational data available at the global scale. For instance, the datasets available to FishMIP are mostly derived from global catch reconstructions (Watson & Tidd, 2018). Recently, a fisheries-independent dataset of biomass from bottom trawl surveys became available, but it only covers coastal regions in the Northern Hemisphere, and authors suggest that biomass cannot be compared across regions (Maureaud et al., 2023). At the regional scale, in several instances, there is enough data to conduct calibration, but the availability of appropriate optimisation routines can constrain the application of systematic calibration of regional MEMs (Oliveros-Ramos & Shin, 2016). To address these issues, FishMIP aims to develop standardised datasets to evaluate historical model simulations (Blanchard et al., 2024), standardised methodological frameworks for model skill evaluation, novel approaches to exploring how best to constrain projections (Novaglio et al., this issue), and novel lightweight approaches to systematically execute and assess MEMs (Steenbeek et al., 2024). These actions will support the development of model benchmarks and tools (Collier et al., 2018; Fu et al., 2022) and ultimately lead to improved ecosystem models. This implementation framework represents one of these actions by standardising model forcings and observational datasets and ultimately reducing model parameterization uncertainty (Blanchard et al., 2024).

5.3 Global-regional model intercomparison

The FishMIP 3a protocol permits the use of standardised fishing effort for global and regional models. While regional ecosystem modellers may find the global effort forcing less precise for their regions compared to local data due to factors such as the taxonomic resolution of the forcing (functional groups instead of species) and system specific variation in catch or effort reporting not captured in the global reconstructions, the standardised fishing effort allows modellers to conduct systematic comparisons between global and regional MEMs. This is one of the main challenges for FishMIP and a priority area for future work, as it will enable us to determine if the projections from regional MEMs are similar or different to those from global MEMs and the likely causes for these differences (Eddy et al., this issue; Novaglio et al., 2024). Fostering these comparisons is especially important for regional impact assessments in data-limited areas, as they will provide insights into whether projections from global MEMs can be used for regional purposes.

5.4 Insights from using the global fishing effort on regional MEMs

Poor agreement was found between the historical trends of the global and regional fishing efforts for some species, e.g., cod in the Baltic Sea and anchovy in the southern Benguela ecosystem models. This is likely explained by the functional group resolution of global effort data, compared to regional resolution, which was to the species level. Thus, in several instances, one fleet can target different species within the same functional group. For example, both anchovy and sardine were included in the 'pelagic<30cm' functional group targeted by the purse seine fleet in the southern Benguela model. Similarly, in the Baltic Sea model, two species were included in the same functional group and fleet. The level of taxonomic resolution (e.g., functional group), therefore, results in the same temporal variability in effort being applied to the different species within a functional group and gear. This is, however, not always the case for species targeted under the same fleet. The global effort data can thus be less representative for some species within the same functional group, and this could explain why anchovy harvest proportions showed a poor correlation with the global effort estimates, while an acceptable correlation was observed for sardine for the southern Benguela model.

The protocol thus advises modellers to first evaluate how regional observations compare to global data, and the applicability of the latter for a particular region. For instance, the sensitivity analysis presented in step 5 will allow us to determine the impacts of using global vs regional forcings on regional MEM outputs and whether the differences between the effort time series are sufficiently large to impact model outputs and the extent of the impact. We acknowledge that if the differences in trends and magnitudes between the datasets are considerable, it may not be productive for regional modellers to recalibrate their MEMs to the global fishing efforts, which are considered less appropriate than the regional ones. If recalibration cannot be carried out, we still hope modellers will submit their runs and compare them to the outputs of their baseline calibrated runs and regional observations. The latter will help identify areas for improvement and refinement of both global and regional MEMs, and global datasets (e.g. effort data) that are regularly used for other reasons in fisheries and anthropogenic impact assessments. Moreover, it will ultimately contribute to the improvements of MEMs within FishMIP (Heneghan et al., 2021), which are often also used for other purposes, the rigour of which would also benefit from any MEM improvements. Lastly, it will also contribute to efforts by the FishMIP community to include an evaluation approach into the MEM protocol (Blanchard et al., 2024) that could also be used regionally. All of these advances move the entire MEM community more clearly toward best practice standards that could be applied to any MEM at any scale in all project work (Planque et al., 2022; Steenbeek et al., 2021).

5.5 Next steps

Given the large amounts of climate and fishing effort data used for this protocol, the Regional Climate Forcing Data Explorer shiny app is a significant step forward in simplifying the processing of these forcings as it performs some of the common steps (e.g., extraction and subsetting) followed in data processing. Moreover, several R and Python scripts that supplement the data processing and analyses performed in this study are publicly available in the FishMIP GitHub repository to ensure the replicability of the process. In the near future, the shiny app will also integrate the global effort data for the different participating regional MEM areas to further simplify the analysis of forcings and foster the application of this workflow for comparisons across regions and global-regional comparisons.

Another area that requires further attention is the use of a harmonised downscaling approach. While this was an area that needed attention for only specific models in the past, it has become one of the focus areas for future work in FishMIP due to the importance of using highly resolved projections for regional climate-impact assessment and other management applications (Pozo Buil et al., 2021).

6. Conclusions

To date, a range of different methods have been used to process and implement climate forcings in regional MEMs participating in FishMIP, with the decision on the methods used lying with the ecosystem modellers. Moreover, the diversity of approaches to implementing climate impacts on MEMs can limit the ability of researchers to replicate the process and compare and analyse MEM ensemble outputs. To address this concern, we developed a workflow that standardises the analysis of climate and fishing forcings, with a focus on global-regional and regional model intercomparisons. The development of this framework is particularly timely, given the increasing number of regional modellers joining FishMIP and the need to systematically evaluate the impacts of climate change worldwide.

While this workflow is designed for model intercomparisons under FishMIP, it may also be adapted to other climate model-MEM linkages. This is particularly important given that projections under climate change are becoming standard expectations in many jurisdictions as the influence of climate change on marine ecosystems matches or exceeds that of fishing (e.g. Fulton et al., 2024). The steps identified in Figure 2 can be generalised to:

1) Identify climate variables needed for the MEM
2) Develop shapefiles of MEM region to extract variables from climate models
3) Aggregate climate variables for non-spatial models
4) Apply downscaling, if needed, for spatial models
5) Apply appropriate fishing effort
6) Calibrate MEM
7) Set up MEM experimental or scenario runs
8) Perform quality checks

If a regional modeler has an application that requires use of different forcing datasets (e.g., use of more regionally specific fishing effort data than what is available in the global fishing data set), then the user can apply those data as needed. However, as a check to their climate-MEM set-up, they can use the FishMIP forcing data sets and perform the FishMIP quality check as a first pass. The inclusion of these regional simulations in FishMIP will facilitate a broader intercomparison and wider understanding of climate impacts on fishing ecosystems globally. The user could then apply their local forcing data for their final application. Substituting forcing data sets enables the user to test the sensitivity of their climate-MEM to different drivers.

The workflow presented here provides a flexible approach to setting up regional MEMs for hindcasts or projections under different climate and fishing scenarios. This workflow is adaptable to different types of regional MEMs, including those that are aspatial or spatial and fully-depth resolved, and those that include one or several fishing fleets. Despite some limitations in the global effort data, the results shown here support its use in regional MEMs, especially for areas with limited fishing information. It is expected that regional models conduct the simulations as described in this protocol to evaluate differences in MEM outputs when using global vs regional sources, provide recommendations for improving global-regional comparisons, and detect drivers of past change in a standardised manner.

DATA AVAILABILITY

The R scripts used to execute the analyses in the paper can be found at: https://github.com/Fish-MIP/FishMIP_regions, https://github.com/pwoodworth-jefcoats/therMizer-FishMIP-2022-HI/blob/main/ClimateForcing/Temperature/Prep_TempRealms_therMizer.Rmd, https://data.isimip.org/, https://rstudio.global-ecosystem-model.cloud.edu.au/shiny/FishMIP_Input_Explorer/, http://portal.sf.utas.edu.au/thredds/catalog/gem/fishmip/catalog.html, https://github.com/Fish-MIP/Regional_MEM_Model_Templates, https://github.com/Fish-MIP/FishMIP_Input_Explorer/blob/main/data_wrangling/regional_data_extractions_DKRZ.py

References

Audzijonyte, A., Pethybridge, H., Porobic, J., Gorton, R., Kaplan, I., & Fulton, E. A. (2019). Atlantis: A spatially explicit end-to-end marine ecosystem model with dynamically integrated physics, ecology and socio-economic modules. *Methods in Ecology and Evolution*, *10*(10), 1814–1819. https://doi.org/10.1111/2041-210X.13272

Bahlburg, D., Thorpe, S. E., Meyer, B., Berger, U., & Murphy, E. J. (2023). An intercomparison of models predicting growth of Antarctic krill (Euphausia superba): The importance of recognizing model specificity. *PLOS ONE*, *18*(7), e0286036. https://doi.org/10.1371/journal.pone.0286036

Blanchard, J. L., Novaglio, C., Maury, O., Harrison, C. S., Petrik, C. M., Arcos, L. D. F., Ortega-Cisneros, K., Bryndum-Buchholz, A., Eddy, T., Heneghan, R., Roberts, K. E., Schewe, J., Bianchi, D., Guiet, J., Denderen, D. V., Palacios-Abrantes, J., Liu, X., Stock, C. A. A., Rousseau, Y., … Tittensor, D. (2024). *Detecting, attributing, and projecting global marine ecosystem and fisheries change: FishMIP 2.0*. https://doi.org/10.22541/essoar.170594183.33534487/v1

Bryndum-Buchholz, A., Tittensor, D. P., Blanchard, J. L., Cheung, W. W. L., Coll, M., Galbraith, E. D., Jennings, S., Maury, O., & Lotze, H. K. (2019). Twenty-first-century climate change impacts on marine animal biomass and ecosystem structure across ocean basins. *Global Change Biology*, *25*(2), 459–472. https://doi.org/10.1111/gcb.14512

Casanueva, A., Herrera, S., Iturbide, M., Lange, S., Jury, M., Dosio, A., Maraun, D., & Gutiérrez, J. M. (2020). Testing bias adjustment methods for regional climate change applications under observational uncertainty and resolution mismatch. *Atmospheric Science Letters*, *21*(7), e978. https://doi.org/10.1002/asl.978

Cheung, W. W. L., Lam, V. W. Y., Sarmiento, J. L., Kearney, K., Watson, R. E. G., Zeller, D., & Pauly, D. (2010). Large-scale redistribution of maximum fisheries catch potential in the global ocean under climate change. *Global Change Biology*, *16*(1), 24–35. https://doi.org/10.1111/j.1365-2486.2009.01995.x

Christensen, V., Coll, M., Steenbeek, J., Buszowski, J., Chagaris, D., & Walters, C. J. (2014). Representing Variable Habitat Quality in a Spatial Food Web Model. *Ecosystems*, *17*(8), 1397–1412. https://doi.org/10.1007/s10021-014-9803-3

Christensen, V., & Walters, C. J. (2004). Ecopath with Ecosim: Methods, capabilities and limitations. *Ecological Modelling*, *172*(2), 109–139. https://doi.org/10.1016/j.ecolmodel.2003.09.003

Cisneros-Montemayor, A. M., Zetina-Rejón, M. J., Espinosa-Romero, M. J., Cisneros-Mata, M. A., Singh, G. G., & Fernández-Rivera Melo, F. J. (2020). Evaluating ecosystem impacts of data-limited artisanal fisheries through ecosystem modelling and traditional fisher knowledge. *Ocean & Coastal Management*, *195*, 105291. https://doi.org/10.1016/j.ocecoaman.2020.105291

Coll, M., Shannon, L. J., Kleisner, K. M., Juan-Jordá, M. J., Bundy, A., Akoglu, A. G., Banaru, D., Boldt, J. L., Borges, M. F., Cook, A., Diallo, I., Fu, C., Fox, C., Gascuel, D., Gurney, L. J., Hattab, T., Heymans, J. J., Jouffre, D., Knight, B. R., … Shin, Y. J. (2016). Ecological indicators to capture the effects of fishing on biodiversity and conservation status of marine ecosystems. *Ecological Indicators*, *60*, 947–962. http://dx.doi.org/10.1016/j.ecolind.2015.08.048

Collier, N., Hoffman, F. M., Lawrence, D. M., Keppel-Aleks, G., Koven, C. D., Riley, W. J., Mu, M., & Randerson, J. T. (2018). The International Land Model Benchmarking (ILAMB) System: Design, Theory, and Implementation. *Journal of Advances in Modeling Earth Systems*, *10*(11), 2731–2754. https://doi.org/10.1029/2018MS001354

de Moor, C. (2021). *Sardine projections based on constant catch scenarios. University of Cape Town. Report. Https://doi.org/10.25375/uct.13635128.v1* (p. 36).

de Mutsert, K., Steenbeek, J., Lewis, K., Buszowski, J., Cowan, J. H., & Christensen, V. (2016). Exploring effects of hypoxia on fish and fisheries in the northern Gulf of Mexico using a dynamic spatially explicit ecosystem model. *Ecological Modelling*, *331*, 142–150. https://doi.org/10.1016/J.ECOLMODEL.2015.10.013

Delius, G., Scott, F., Blanchard, J., & Andersen, K. (2023). *Mizer: Dynamic Multi-Species Size Spectrum Modelling. R package version 2.5.0.* https://github.com/sizespectrum/mizer, https://sizespectrum.org/mizer/

DFFE. (2023). *Status of the South African Marine Fishery Resources 2023*. Department of Forestry, Fisheries and the Environment. Cape Town: DFFE, South Africa.

Eddy, TD., Heneghan, RF., Bryndum-Buchholz, A., et al. Global and regional marine ecosystem model climate change projections reveal key uncertainties. Submitted to Earths Future.

Espinoza-Morriberon, D., Echevin, V., Tam, J., Ledesma, J., Oliveros-Ramos, R., Ramos, J., & Romero, C. Y. (2016). Biogeochemical validation of an interannual simulation of the ROMS-PISCES coupled model in the Southeast Pacific. *Revista Peruana de Biología*, *23*(2), 159–168. https://doi.org/10.15381/rpb.v23i2.12427

Frieler, K., Volkholz, J., Lange, S., Schewe, J., Mengel, M., del Rocío Rivas López, M., Otto, C., Reyer, C. P. O., Karger, D. N., Malle, J. T., Treu, S., Menz, C., Blanchard, J. L., Harrison, C. S., Petrik, C. M., Eddy, T. D., Ortega-Cisneros, K., Novaglio, C., Rousseau, Y., … Bechtold, M. (2024). Scenario setup and forcing data for impact model evaluation and impact attribution within the third round of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP3a). *Geoscientific Model Development*, *17*(1), 1–51. https://doi.org/10.5194/gmd-17-1-2024

Fu, C., Xu, Y., Bundy, A., Grüss, A., Coll, M., Heymans, J. J., Fulton, E. A., Shannon, L., Halouani, G., & Velez, L. (2019). Making ecological indicators management ready: Assessing the specificity, sensitivity, and threshold response of ecological indicators. *Ecological Indicators*, *105*, 16–28.

Fu, W., Moore, J. K., Primeau, F., Collier, N., Ogunro, O. O., Hoffman, F. M., & Randerson, J. T. (2022). Evaluation of Ocean Biogeochemistry and Carbon Cycling in CMIP Earth System Models With the International Ocean Model Benchmarking (IOMB) Software System. *Journal of Geophysical Research: Oceans*, *127*(10), e2022JC018965. https://doi.org/10.1029/2022JC018965

Fulton, E. A. (2011). Interesting times: Winners, losers, and system shifts under climate change around Australia. *ICES Journal of Marine Science: Journal Du Conseil*, *68*(6), 1329–1342. https://doi.org/10.1093/icesjms/fsr032

Fulton, E. A., Mazloumi, N., Puckeridge, A., & Hanamseth, R. (2024). Modelling perspective on the climate footprint in south east Australian marine waters and its fisheries. *ICES Journal of Marine Science*, *81*(1), 130–144. https://doi.org/10.1093/icesjms/fsad185

Garcia, H., Boyer, T., Baranova, O., Locarnini, R., Mishonov, A., Grodsky, A., Paver, C., Weathers, K., Smolyar, I., Reagan, J., Seidow, D., & Zweng, M. (2019). *World Ocean Atlas 2018: Product Documentation. A. Mishonov, Technical Editor.* https://www.ncei.noaa.gov/data/oceans/woa/WOA18/DOC/woa18documentation.pdf

Heneghan, R. F., Galbraith, E., Blanchard, J. L., Harrison, C., Barrier, N., Bulman, C., Cheung, W., Coll, M., Eddy, T. D., Erauskin-Extramiana, M., Everett, J. D., Fernandes-Salvador, J. A., Gascuel, D., Guiet, J., Maury, O., Palacios-Abrantes, J., Petrik, C. M., du Pontavice, H., Richardson, A. J., … Tittensor, D. P. (2021). Disentangling diverse responses to climate change among global marine ecosystem models. *Progress in Oceanography*, *198*, 102659. https://doi.org/10.1016/j.pocean.2021.102659

Heymans, J. J., Link, J. S., Mackinson, S., Steenbeek, J., Walters, C., Christensen, V., Coll, M., Link, J. S., Mackinson, S., Steenbeek, J., Walters, C., & Christensen, V. (2016). Best practice in Ecopath with Ecosim food-web models for ecosystem-based management. *Ecological Modelling*, *331*, 173. http://dx.doi.org/10.1016/j.ecolmodel.2015.12.007

Hipsey, M. R., Gal, G., Arhonditsis, G. B., Carey, C. C., Elliott, J. A., Frassl, M. A., Janse, J. H., de Mora, L., & Robson, B. J. (2020). A system of metrics for the assessment and improvement of aquatic ecosystem models. *Environmental Modelling & Software*, *128*, 104697. https://doi.org/10.1016/j.envsoft.2020.104697

IPCC. (2023). *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press. https://doi.org/10.1017/9781009157896

Kristiansen, T., Butenschön, M., & Peck, M. A. (2024). Statistically downscaled CMIP6 ocean variables for European waters. *Scientific Reports*, *14*(1), 1209. https://doi.org/10.1038/s41598-024-51160-1

Lange, S. (2019). Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0). *Geoscientific Model Development*, *12*(7), 3055–3070. https://doi.org/10.5194/gmd-12-3055-2019

Lindmark, M., Audzijonyte, A., Blanchard, J. L., & Gårdmark, A. (2022). Temperature impacts on fish physiology and resource abundance lead to faster growth but smaller fish sizes and yields under warming. *Global Change Biology*, *28*(21), 6239–6253. https://doi.org/10.1111/gcb.16341

Liu, X., Stock, C. A., Dunne, J. P., Lee, M., Shevliakova, E., Malyshev, S., & Milly, P. C. D. (2021). Simulated Global Coastal Ecosystem Responses to a Half-Century Increase in River Nitrogen Loads. *Geophysical Research Letters*, *48*(17), e2021GL094367. https://doi.org/10.1029/2021GL094367

Locarnini, R., Mishonov, A., Baranova, O., Boyer, T., Zweng, M., Garcia, H., Reagan, J., Seidow, D., Weathers, K., Paver, C., & Smolyar, I. (2018). *World Ocean Atlas 2018, Volume 1: Temperature. A. Mishonov Technical Ed.; NOAA Atlas NESDIS 81* (p. 52). https://www.ncei.noaa.gov/data/oceans/woa/WOA18/DOC/woa18documentation.pdf

Lotze, H. K., Tittensor, D. P., Bryndum-Buchholz, A., Eddy, T. D., Cheung, W. W. L., Galbraith, E. D., Barange, M., Barrier, N., Bianchi, D., Blanchard, J. L., Bopp, L., Büchner, M., Bulman, C. M., Carozza, D. A., Christensen, V., Coll, M., Dunne, J. P., Fulton, E. A., Jennings, S., … Worm, B. (2019). Global ensemble projections reveal trophic amplification of ocean biomass declines with climate change. *Proceedings of the National Academy of Sciences of the United States of America*, 201900194. https://doi.org/10.1073/pnas.1900194116

Luján, C., Oliveros-Ramos, R., Barrier, N., Leadley, P., & Shin, Y.-J. (2024). Key species and indicators revealed by an uncertainty analysis of the marine ecosystem model OSMOSE. *Marine Ecology Progress Series*, *SPF2*. https://doi.org/10.3354/meps14465

Luo, Y. Q., Randerson, J. T., Abramowitz, G., Bacour, C., Blyth, E., Carvalhais, N., Ciais, P., Dalmonech, D., Fisher, J. B., Fisher, R., Friedlingstein, P., Hibbard, K., Hoffman, F., Huntzinger, D., Jones, C. D., Koven, C., Lawrence, D., Li, D. J., Mahecha, M., … Zhou, X. H. (2012). A framework for benchmarking land models. *Biogeosciences*, *9*(10), 3857–3874. https://doi.org/10.5194/bg-9-3857-2012

Mackinson, S., Platts, M., Garcia, C., & Lynam, C. (2018). Evaluating the fishery and ecological consequences of the proposed North Sea multi-annual plan. *PLOS ONE*, *13*(1), e0190015. https://doi.org/10.1371/journal.pone.0190015

Marshall, K. N., Kaplan, I. C., Hodgson, E. E., Hermann, A., Busch, D. S., McElhany, P., Essington, T. E., Harvey, C. J., & Fulton, E. A. (2017). Risks of ocean acidification in the California Current food web and fisheries: Ecosystem model projections. *Global Change Biology*, *23*(4), 1525–1539. https://doi.org/10.1111/gcb.13594

Maureaud, A., Kitchel, Z., Fredston, A., Guralnick, R., Abrantes, J. P., Palomares, D., Pinsky, M., Shackell, N., Thorson, J., Merigot, B., & Consortium, F. (2023). *FISHGLOB: A collaborative infrastructure for marine science and management*. OSF. https://doi.org/10.31219/osf.io/mh46b

Maury O., D. P. Tittensor, T. D. Eddy, E. H. Allison, T. Bahri, N. Barrier, L. Campling, W. W. L. Cheung, K. Frieler, E. A. Fulton, P. Guillotreau, R F. Heneghan, V. W. Y. Lam, D. Leclère, M. Lengaigne, H. Lotze-Campen, C. Novaglio, K. Ortega-Cisneros, J. Schewe, Y.-J. Shin, H. Sloterdijk, D. Squires, U. R. Sumaila, A. N. Tidd, B. van Ruijven and J. Blanchard, submitted. The Ocean System Pathways (OSPs): a new scenario and simulation framework to investigate the future of the world fisheries. The Past and Future of the Fisheries and Marine Ecosystem Model Intercomparison Project. Submitted to Earth's Future.

Melsom, A., Lien, V. S., & Budgell, W. P. (2009). Using the Regional Ocean Modeling System (ROMS) to improve the ocean circulation from a GCM 20th century simulation. *Ocean Dynamics*, *59*(6), 969–981. https://doi.org/10.1007/s10236-009-0222-5

Murphy, K., Fierros-Arcos, D., Rohr, T., Green, D., Novaglio, C., et al. Developing a Southern Ocean Marine Ecosystem Model Ensemble To Assess Climate Risks and Uncertainties. Submitted to Earth's Future.

Novaglio, C., Bryndum-Buchholz, A., Tittensor, D., Eddy, T., Lotze, H. K., Harrison, C. S., Heneghan, R., Maury, O., Ortega-Cisneros, K., Petrik, C. M., Roberts, K. E., & Blanchard, J. L. (2024). *The Past and Future of the Fisheries and Marine Ecosystem Model Intercomparison Project*. https://doi.org/10.22541/essoar.170542252.20348236/v1

Ogunro, O. O., Elliott, S. M., Wingenter, O. W., Deal, C., Fu, W., Collier, N., & Hoffman, F. M. (2018). Evaluating Uncertainties in Marine Biogeochemical Models: Benchmarking Aerosol Precursors. *Atmosphere*, *9*(5), Article 5. https://doi.org/10.3390/atmos9050184

Oliveros-Ramos, R., & Shin, Y.-J. (2016). *Calibrar: An R package for fitting complex ecological models* (arXiv:1603.03141). arXiv. https://doi.org/10.48550/arXiv.1603.03141

Oliveros-Ramos, R., Shin, Y.-J., Gutiérrez, D., & Trenkel, V. (2023). *A multi-model selection approach for statistical downscaling and bias correction of Earth System Model outputs for regional impact applications*. https://doi.org/10.22541/essoar.167810427.75944849/v1

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., & Sanderson, B. M. (2016). The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, *9*(9), 3461–3482. https://doi.org/10.5194/gmd-9-3461-2016

Ortega-Cisneros, K., Cochrane, K., & Fulton, E. A. E. A. (2017). An Atlantis model of the southern Benguela upwelling system: Validation, sensitivity analysis and insights into ecosystem functioning. *Ecological Modelling*, *355*, 49–63. https://doi.org/10.1016/j.ecolmodel.2017.04.009

Ortega-Cisneros, K., Cochrane, K. L., Fulton, E. A., Gorton, R., & Popova, E. (2018). Evaluating the effects of climate change in the southern Benguela upwelling system using the Atlantis modelling framework. *Fisheries Oceanography*, *27*, 489–503. https://doi.org/10.1111/fog.12268

Ortega-Cisneros, K., Shannon, L., Cochrane, K., Fulton, E. A., & Shin, Y.-J. (2018). Evaluating the specificity of ecosystem indicators to fishing in a changing environment: A model comparison study for the southern Benguela ecosystem. *Ecological Indicators*, *95*, 85–98. https://doi.org/10.1016/J.ECOLIND.2018.07.021

Palacios-Abrantes, J., Frölicher, T. L., Reygondeau, G., Sumaila, U. R., Tagliabue, A., Wabnitz, C. C. C., & Cheung, W. W. L. (2022). Timing and magnitude of climate-driven range shifts in transboundary fish stocks challenge their management. *Global Change Biology*, *28*(7), 2312–2326. https://doi.org/10.1111/gcb.16058

Pauly, D., Zeller, D., & Palomares, M. (2020). *Sea Around Us Concepts, Design and Data (seaaroundus.org).*

Pethybridge, H. R., Fulton, E. A., Hobday, A. J., Blanchard, J., Bulman, C. M., Butler, I. R., Cheung, W. W. L., Dutra, L. X. C., Gorton, R., Hutton, T., Matear, R., Lozano-Montes, H., Plagányi, E. E., Villanueva, C., & Zhang, X. (2020). Contrasting Futures for Australia's Fisheries Stocks Under IPCC RCP8.5 Emissions – A Multi-Ecosystem Model Approach. *Frontiers in Marine Science*, *7*. https://www.frontiersin.org/articles/10.3389/fmars.2020.577964

Plagányi, É. E., Punt, A. E., Hillary, R., Morello, E. B., Thébaud, O., Hutton, T., Pillans, R. D., Thorson, J. T., Fulton, E. A., Smith, A. D. M., Smith, F., Bayliss, P., Haywood, M., Lyne, V., & Rothlisberg, P. C. (2014). Multispecies fisheries management and conservation: Tactical applications using models of intermediate complexity. *Fish and Fisheries*, *15*(1), 1–22. https://doi.org/10.1111/j.1467-2979.2012.00488.x

Pörtner, H.-O., Scholes, R. J., Agard, J., Archer, E., Arneth, A., Bai, X., Barnes, D., Burrows, M., Chan, L., Cheung, W. L. (William) ., Diamond, S., Donatti, C., Duarte, C., Eisenhauer, N., Foden, W., Gasalla, M. A., Handa, C., Hickler, T., Hoegh-Guldberg, O., … Ngo, H. (2021). Scientific outcome of the IPBES-IPCC co-sponsored workshop on biodiversity and climate change (Version 5). Zenodo. https://doi.org/10.5281/zenodo.5101125

Pozo Buil, M., Fiechter, J., Jacox, M. G., Bograd, S. J., & Alexander, M. A. (2023). Evaluation of Different Bias Correction Methods for Dynamical Downscaled Future Projections of the California Current Upwelling System. *Earth and Space Science*, *10*(12), e2023EA003121. https://doi.org/10.1029/2023EA003121

Pozo Buil, M., Jacox, M. G., Fiechter, J., Alexander, M. A., Bograd, S. J., Curchitser, E. N., Edwards, C. A., Rykaczewski, R. R., & Stock, C. A. (2021). A Dynamically Downscaled Ensemble of Future Projections for the California Current System. *Frontiers in Marine Science*, *8*. https://doi.org/10.3389/fmars.2021.612874

Reed, J., Shannon, L., Velez, L., Akoglu, E., Bundy, A., Coll, M., Fu, C., Fulton, E. A., Grüss, A., Halouani, G., Heymans, J. J., Houle, J. E., John, E., Le Loc'h, F., Salihoglu, B., Verley, P., & Shin, Y.-J. (2016). Ecosystem indicators—Accounting for variability in species' trophic levels. *ICES Journal of Marine Science: Journal Du Conseil*. https://doi.org/10.1093/icesjms/fsw150

Reum, J. C. P., Woodworth-Jefcoats, P., Novaglio, C., Forestier, R., Audzijonyte, A., Gårdmark, A., Lindmark, M., & Blanchard, J. L. (2024). Temperature-Dependence Assumptions Drive Projected Responses of Diverse Size-Based Food Webs to Warming. *Earth's Future*, *12*(3), e2023EF003852. https://doi.org/10.1029/2023EF003852

Rocklöv, J., Huber, V., Bowen, K., & Paul, R. (2021). Taking globally consistent health impact projections to the next level. *The Lancet Planetary Health*, *5*(7), e487–e493. https://doi.org/10.1016/S2542-5196(21)00171-6

Rogers, A. D., Frinault, B. a. V., Barnes, D. K. A., Bindoff, N. L., Downie, R., Ducklow, H. W., Friedlaender, A. S., Hart, T., Hill, S. L., Hofmann, E. E., Linse, K., McMahon, C. R., Murphy, E. J., Pakhomov, E. A., Reygondeau, G., Staniland, I. J., Wolf-Gladrow, D. A., & Wright, R. M. (2020). Antarctic Futures: An Assessment of Climate-Driven

Changes in Ecosystem Structure, Function, and Service Provisioning in the Southern Ocean. *Annual Review of Marine Science*, *12*(Volume 12, 2020), 87–120. https://doi.org/10.1146/annurev-marine-010419-011028

Rosenzweig, C., Jones, J. W., Hatfield, J. L., Ruane, A. C., Boote, K. J., Thorburn, P., Antle, J. M., Nelson, G. C., Porter, C., Janssen, S., Asseng, S., Basso, B., Ewert, F., Wallach, D., Baigorria, G., & Winter, J. M. (2013). The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies. *Agricultural and Forest Meteorology*, *170*, 166–182. https://doi.org/10.1016/j.agrformet.2012.09.011

Rousseau, Y., Blanchard, J. L., Novaglio, C., Pinnell, K. A., Tittensor, D. P., Watson, R. A., & Ye, Y. (2024). A database of mapped global fishing activity 1950–2017. *Scientific Data*, *11*(1), 48. https://doi.org/10.1038/s41597-023-02824-6

Rovellini, A., Punt, A. E., Bryan, M. D., Kaplan, I. C., Dorn, M. W., Aydin, K., Fulton, E. A., Alglave, B., Baker, M. R., Carroll, G., Ferriss, B. E., Haltuch, M. A., Hayes, A. L., Hermann, A. J., Hernvann, P.-Y., Holsman, K. K., Liu, O. R., McHuron, E., Morzaria-Luna, H. N., … Weise, M. T. (2024). Linking climate stressors to ecological processes in ecosystem models, with a case study from the Gulf of Alaska. *ICES Journal of Marine Science*, fsae002. https://doi.org/10.1093/icesjms/fsae002

Ruzicka, J. J., Brink, K. H., Gifford, D. J., & Bahr, F. (2016). A physically coupled end-to-end model platform for coastal ecosystems: Simulating the effects of climate change and changing upwelling characteristics on the Northern California Current ecosystem. *Ecological Modelling*, *331*, 86–99. https://doi.org/10.1016/j.ecolmodel.2016.01.018

Shin, Y.-J., Bundy, A., Shannon, L. J., Simier, M., Coll, M., Fulton, E. A., Link, J. S., Jouffre, D., Ojaveer, H., Mackinson, S., Heymans, J. J., & Raid, T. (2010). Can simple be useful and reliable? Using ecological indicators to represent and compare the states of marine ecosystems. *ICES Journal of Marine Science: Journal Du Conseil*, *67*(4), 717–731. https://doi.org/10.1093/icesjms/fsp287

Shin, Y.-J., & Cury, P. (2001). Exploring fish community dynamics through size-dependent trophic interactions using a spatialized individual-based model. *Aquatic Living Resources*, *14*(2), 65–80. http://dx.doi.org/10.1016/S0990-7440(01)01106-8

Shin, Y.-J., Houle, J. E., Akoglu, E., Blanchard, J. L., Bundy, A., Coll, M., Demarcq, H., Fu, C., Fulton, E. A., Heymans, J. J., Salihoglu, B., Shannon, L., Sporcic, M., & Velez, L. (2018). The specificity of marine ecological indicators to fishing in the face of environmental change: A multi-model evaluation. *Ecological Indicators*, *89*, 317–326. https://doi.org/10.1016/j.ecolind.2018.01.010

Shin, Y.-J., Shannon, L. J., Bundy, A., Coll, M., Aydin, K., Bez, N., Blanchard, J. L., Borges, M. de F., Diallo, I., Diaz, E., Heymans, J. J., Hill, L., Johannesen, E., Jouffre, D., Kifani, S., Labrosse, P., Link, J. S., Mackinson, S., Masski, H., … Cury, P. M. (2010). Using indicators for evaluating, comparing, and communicating the ecological status of exploited marine ecosystems. 2. Setting the scene. *ICES Journal of Marine Science: Journal Du Conseil*, *67*(4), 692–716. https://doi.org/10.1093/icesjms/fsp294

Skogen, M. D., Hjøllo, S. S., Sandø, A. B., & Tjiputra, J. (2018). Future ecosystem changes in the Northeast Atlantic: A comparison between a global and a regional model system. *ICES Journal of Marine Science*, *75*(7), 2355–2369. https://doi.org/10.1093/icesjms/fsy088

Steenbeek, J., Ortega, P., Bernardello, R., Christensen, V., Coll, M., Exarchou, E., Fuster-Alonso, A., Heneghan, R., Julià Melis, L., Pennino, M. G., Rivas, D., & Keenlyside, N. (2024). Making Ecosystem Modeling Operational–A Novel Distributed Execution Framework to Systematically Explore Ecological Responses to Divergent Climate Trajectories. *Earth's Future*, *12*(3), e2023EF004295.

https://doi.org/10.1029/2023EF004295
Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*, *93*(4), 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1
Tittensor, D. P., Eddy, T. D., Lotze, H. K., Galbraith, E. D., Cheung, W., Barange, M., Blanchard, J. L., Bopp, L., Bryndum-Buchholz, A., Büchner, M., Bulman, C., Carozza, D. A., Christensen, V., Coll, M., Dunne, J. P., Fernandes, J. A., Fulton, E. A., Hobday, A. J., Huber, V., … Walker, N. D. (2018). A protocol for the intercomparison of marine fishery and ecosystem models: Fish-MIP v1.0. *Geoscientific Model Development*, *11*(4), 1421–1442. https://doi.org/10.5194/gmd-11-1421-2018
Tittensor, D. P., Novaglio, C., Harrison, C. S., Heneghan, R. F., Barrier, N., Bianchi, D., Bopp, L., Bryndum-Buchholz, A., Britten, G. L., Büchner, M., Cheung, W. W. L., Christensen, V., Coll, M., Dunne, J. P., Eddy, T. D., Everett, J. D., Fernandes-Salvador, J. A., Fulton, E. A., Galbraith, E. D., … Blanchard, J. L. (2021). Next-generation ensemble projections reveal higher climate risks for marine ecosystems. *Nature Climate Change*. https://doi.org/10.1038/s41558-021-01173-9
Tsujino, H., Urakawa, S., Nakano, H., Small, R. J., Kim, W. M., Yeager, S. G., Danabasoglu, G., Suzuki, T., Bamber, J. L., Bentsen, M., Böning, C. W., Bozec, A., Chassignet, E. P., Curchitser, E., Boeira Dias, F., Durack, P. J., Griffies, S. M., Harada, Y., Ilicak, M., … Yamazaki, D. (2018). JRA-55 based surface dataset for driving ocean–sea-ice models (JRA55-do). *Ocean Modelling*, *130*, 79–139. https://doi.org/10.1016/j.ocemod.2018.07.002
Tulloch, V. J. D., Plagányi, É. E., Brown, C., Richardson, A. J., & Matear, R. (2019). Future recovery of baleen whales is imperiled by climate change. *Global Change Biology*, *25*(4), 1263–1281. https://doi.org/10.1111/gcb.14573
Watson, R. A., & Tidd, A. (2018). Mapping nearly a century and a half of global marine fishing: 1869–2015. *Marine Policy*, *93*, 171–177. https://doi.org/10.1016/j.marpol.2018.04.023
Woodworth-Jefcoats, P. A., Blanchard, J. L., & Drazen, J. C. (2019). Relative Impacts of Simultaneous Stressors on a Pelagic Marine Ecosystem. *Frontiers in Marine Science*, *6*. https://www.frontiersin.org/articles/10.3389/fmars.2019.00383

Table 1. A description of climate forcings used by, and how process effects are represented in, the FishMIP ecosystem modelling types. The optional forcing column highlights variables that may be required by some implementations of the model type. The key forcings column represents those for which the climate input forcings described in the workflow are preferred over time series drawn from other sources or defined by default within the model. Adapted from Tittensor et al., 2018.

Ecosystem model name	Spatial and temporal scale and vertical resolutions	Key forcing variables used	Optional forcing variables used	Implementation of temperature effects/processes	Implementation of primary production / Plankton biomass
Composite (hybrid) models – including multiple model formulations in system representation					
Atlantis	3-D spatial polygons matched to biophysical features; vertically resolved using "slab" layers (with finer layers and the surface and thicker at depth). Timestep is flexible, typically 6–24 h	Sea Water Potential Temperature (thetao), Sea Water Salinity (so), Sea water X velocity (uo), Sea water Y velocity (vo)	Dissolved oxygen concentration (o_2), pH (pH), Mole Concentration of nutrients (NH, NO, Si and potentially micronutrients), Diatoms (phydiat), Diazotrophs (phydiaz), Picophytoplankton (phypico), Sea ice, irradiance, precipitation, river inflow, changes in sea level, eddy strength	Any model ecological process (e.g. metabolic rates, consumption, growth, mortality, movement/distribution, spawning) and the functional groups as defined by the modeller, as well as all modelled biogeochemical processes	Plankton mole concentration (in N m^{-3}) read in and forcing replaces the emergent phytoplankton biomass estimated for each model. Best done as a weighted average (somewhat similar to data assimilation), to minimise loss of mass conservation. Delta method to correct primary production or plankton biomass can be applied as a relative anomaly to the phytoplankton growth rates (Rovellini et al., 2024)
Models of Intermediate Complexity (MICE)	Flexible, typically running in monthly or yearly time steps. Can be non-spatial or spatial. If spatial, applications are usually of coarse resolution. Spatially resolved in 2-D.	Sea Water Potential Temperature (thetao) averaged over specific depth ranges to represent the preferences of the different functional groups or species included in the model;	Chl-a, Primary Organic Carbon Production by All Types of Phytoplankton (intpp), Mole concentration of Diatoms (phydiat), Diazotrophs (phydiaz), Picophytoplankton (phypico), mesozooplankton (zmeso) and microzooplankton	Different model ecological process (e.g. growth, mortality, movement/ distribution, spawning) and the functional groups as defined by the modeller	Estimating multipliers for carrying capacity and predator-prey interactions (e.g., Tulloch et al., 2019)

		or at the surface, bottom of the water column*	(zmicro), Rainfall, Sea ice		
OSMOSE	Flexible. Typically, resolution of 1/6 and a weekly time step. Spatially resolved in 2-D; the vertical distribution of species is handled through a matrix of accessibility.	Sea Water Potential Temperature (thetao)*, Primary Organic Carbon Production by All Types of Phytoplankton (intpp), Mole Concentration of Diatoms (phydiat), Diazotrophs (phydiaz), Picophytoplankton (phypico), mesozooplankton (zmeso) and microzooplankton (zmicro)	Sea Water Salinity (so), Dissolved oxygen concentration (o2) (e.g., Moullec et al. 2019, Morell et al. 2023).	Species distributions, Maintenance rate, growth, fecundity, starvation mortality. OSMOSE parameterisation relies on species distribution model outputs. Climate forcings must be within the same range as the data used for parameterisation of thermal preferences to avoid species collapses (e.g., out-of-range environmental conditions)	Statistical downscaling and bias-correction to produce plankton biomass consistent with regional biogeochemical model (ROMS-PISCES) (Espinoza-Morriberon et al., 2016).
Trophodynamic models – structured based on species interactions and transfer of energy across trophic levels.					
EcoTran (Coupled physical-trophic model)	2D and 3D implementations. Rectangular polygons of varying size, ~10s-100s km, and 2-6 depth layers of varying thickness, ~10s-100s m. Time-step typically 24 h nearshore, but 3 h in oceanic regions to simulate diel vertical migration	Temperatures within specific depth ranges. Horizontal water velocities. Nutrient (N) input rate or phytoplankton production rate (flexible phytoplankton group definitions).	User-defined changes to consumption rates of individual consumer groups or catch rates by individual fleets. User-defined changes to community composition and food web structure.	Metabolic rate (Q10). Feeding rate of poikilotherms is scaled via dome-shape response representing optimal and sub-optimal/lethal conditions)	When driven via nitrate and ammonium input, primary production is estimated via Michaelis-menten kinetics. Model may also be driven directly with phytoplankton biomass time-series output of a biogeochemical model (in cases where biomass is available but not production, a constant production/biomass ratio is typically assumed to estimate primary production rates).

EwE	Flexible, typically running in monthly time steps. Depth dimension is considered implicitly through food web interactions and habitat preference patterns for Ecopath and Ecosim. Ecospace is spatially resolved in 2-D; the vertical distribution of species is handled through the niche model (Christensen et al., 2014; de Mutsert et al., 2023).	Sea Water Potential Temperature (thetao)*, Primary Organic Carbon Production by All Types of Phytoplankton (intpp)	Sea Water Salinity (so), Dissolved oxygen concentration (o2)	Typically uses forcing and environmental response functions to model temperature effects through changes in assimilation efficiency, adjustment of consumption rates and mortality. In Ecospace, sea water temperature also affects species distributions.	Primary production used as a forcing function influencing the production of plankton size classes. Primary production from ESMs is bias corrected using the delta method (Eddy et al. this issue). This method involves calculating relative values of primary production compared to the model base year. EwE can also ingest primary producer biomass density distributions directly, overriding internal primary production growth dynamics (de Mutsert et al. 2023)
Size-based models – developed from food web, macroecological, and life history theory for exploration of community size spectra					
(ther)Mizer	Non-spatial. Mizer is a multi-species size-structured model, and therMizer allows climate and plankton forcing to be added to Mizer (Delius et al., 2023; Woodworth-Jefcoats et al., 2019).	Sea Water Potential Temperature (thetao)*, Mole Concentration of Diatoms (phydiat), Diazotrophs (phydiaz), Picophytoplankton (phypico), mesozooplankton (zmeso) and microzooplankton (zmicro)		Individual metabolism, maximum consumption, search volume and predation mortality	Use the concentration of vertically integrated plankton size classes to estimate the plankton size spectrum via linear fit across these size classes

1163

Table 2. List of gear and functional group codes.

Gear codes	Functional groups
Dredges	bathydemersal<30cm
Gillnets	bathydemersal30-90cm
Lift_Nets	bathydemersal>=90cm
Lines_Handlines_and_poles	bathypelagic<30cm
Lines_Longlines	bathypelagic30-90cm
Lines_Unspecified	bathypelagic>=90cm
Others_Multiple_Gears	benthopelagic<30cm
Others_Others	benthopelagic30-90cm
Others_Support	benthopelagic>=90cm
Others_Unknown	cephalopods
Pots_and_Traps	demersal<30cm
Seine_Danish_and_Other	demersal30-90cm
Seine_Purse_Seine	demersal>=90cm
Trawl_Midwater_or_Unsp	demersalmollusc
Trawl_Bottom	flatfish<90cm
Falling_Gear	flatfish>=90cm
	krill
	lobsterscrab
	pelagic<30cm
	pelagic30-90cm
	pelagic>=90cm
	rays<90cm
	rays>=90cm
	reef-associated<30cm
	reef-associated30-90cm
	reef-associated>=90cm

	shark<90cm
	shark>=90cm
	Shrimp