Making ecosystem modelling operational - a novel distributed execution framework to systematically explore ecological responses to divergent climate trajectories

Jeroen Gerhard Steenbeek[1], Pablo Ortega[2], Raffaele Bernardello[2], Villy Christensen[3], Marta Coll[4], Eleftheria Exarchou[5], Alba Fuster[6], Ryan Heneghan[7], Laura Julià[6], Maria Grazia Pennino[8], David Rivas Camargo[9], and Noel S Keenlyside[9]

[1]Ecopath International Initiative (EII)
[2]Barcelona Supercomputing Center
[3]UBC Insitute of the Oceans and Fisheries
[4]Institute of Marine Science (ICM-CSIC), Barcelona, Spain
[5]Barcelona Supercomputing Centre
[6]Institute of Marine Science (ICM-CSIC)
[7]University of Queensland
[8]Instituto Español de Oceanografía
[9]Geophysical Institute, University of Bergen

January 18, 2024

Making ecosystem modelling operational - a novel distributed execution framework to systematically explore ecological responses to divergent climate trajectories

Jeroen Steenbeek[1], Pablo Ortega[2], Raffaele Bernardello[2], Villy Christensen[1,3], Marta Coll[1,4], Eleftheria Exarchou[2], Alba Fuster-Alonso[4], Ryan Heneghan[5], Laura Julià Melis[4], Maria Grazia Pennino[6], David Rivas[7,8], Noel Keenlyside[7,9]

[1] Ecopath International Initiative (EII) Research Association, Barcelona, Spain

[2] Barcelona Supercomputing Center (BSC), Barcelona, Spain

[3] UBC Institute for the Oceans and Fisheries, Vancouver BC, Canada

[4] Institute of Marine Science (ICM-CSIC), Barcelona, Spain

[5] School of Mathematical Sciences, Queensland University of Technology, Brisbane, Queensland, Australia

[6] Instituto Español de Oceanografía, Madrid, Spain

[7] Geophysical Institute, University of Bergen and Bjerknes Centre for Climate Research, Norway

[8] Departamento de Oceanografía Biológica, Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE), Ensenada, Mexico

[9] Nansen Environmental and Remote Sensing Center and Bjerknes Centre for Climate Research, Norway

Corresponding author: Jeroen Steenbeek (jeroen@ecopathinternational.org)

Key Points

Most marine ecosystem modellers lack the skills and resources to systematically calibrate, validate and assess the models for uncertainty

Here we present a low-tech and open source run framework to use any computer network as a distributed model execution and assessment system

We use the framework to mass-execute an Earth System (ESM)/Ecosystem Model (MEM) ensemble to assess the ecosystem impact of ESM uncertainty

Abstract

Marine Ecosystem Models (MEMs) are increasingly driven by Earth System Models (ESMs) to better understand marine ecosystem dynamics, and to analyse the effects of alternative management efforts for marine ecosystems under potential scenarios of climate change. However, policy and commercial activities typically occur on seasonal-to-decadal time scales, a time span widely used in the global climate modelling community but where the skill level assessments of MEMs are in their infancy. This is mostly due to technical hurdles that prevent the global MEM community from performing large ensemble simulations with which to undergo systematic skill assessments. Here, we developed a novel distributed execution framework constructed of low-tech and freely available technologies to enable the systematic execution and analysis of linked ESM / MEM prediction ensembles. We apply this framework on the seasonal-to-decadal time scale, and assess how retrospective forecast uncertainty in an ensemble of initialised decadal ESM predictions affects a mechanistic and spatiotemporal explicit global trophodynamic MEM. Our results indicate that ESM internal variability has a relatively low impact on the MEM variability in comparison to the broad assumptions related to reconstructed fisheries. We also observe that the results are also sensitive to the ESM specificities. Our case study warrants further systematic explorations to disentangle the impacts of climate change, fisheries scenarios, MEM internal ecological hypotheses, and ESM variability. Most importantly, our case study demonstrates that a simple and free distributed execution framework has the potential to empower any modelling group with the fundamental capabilities to operationalize marine ecosystem modelling.

Plain Language Summary

Climate change and human activities like fishing are affecting the balance of marine ecosystems and the services they provide. To understand impacts better, scientists use computer models that consider climate, ocean conditions, and ocean life.

To make robust decisions, decision makers need robust science delivered by robust models. This requires running many computer simulations, but the complexity of marine ecosystems models makes this difficult. Typically, only institutions with sufficient financial and technical means can overcome these difficulties, which leaves the majority of marine ecosystem modellers wanting.

Here we introduce a possible solution to overcome the difficulties, using only simple and free technologies to facilitate the systematic execution of complex ecosystem models across networks of computers, to allow any modelling group to perform these exercises.

We demonstrate our solution by running a global marine ecosystem model, EcoOcean, hundreds of times to see how it is affected by variability in ocean conditions. Using available laptops and desktops, we can now complete this task in 30 hours; where prior this modelling task would have taken weeks to complete. It is conceptually simple solutions such as these that may make the process of marine ecosystem modelling easier and more operational around the globe, thus opening the door for scientific management breakthroughs.

1 Introduction

Climate change and anthropogenic activities such as fishing are having far-reaching consequences for the functioning and stability of marine food webs and the ecosystem services that humanity relies on (e.g, Halpern et al., 2019; Pörtner et al., 2014). To better understand such impacts and their consequences for ocean life and ecosystem services, the global ocean science community increasingly deploys modelling systems that incorporate climate, ocean circulation, biochemistry and marine life under multiple stressors (e.g., Stock et al., 2023). Marine Ecosystem Models (MEMs) forced with Earth System Models (ESMs) are such modelling systems, where ESMs represent the fundamental physical, chemical and biological processes governing the evolution of the Earth system and the interactions within its major components (i.e. atmosphere, ocean, cryosphere and land), while MEMs represent mechanistically the non-linear dynamics between marine species and within marine food webs (Steenbeek et al., 2021; Tittensor et al., 2018).

At present, the scientific agenda on future climate change largely focuses on the decadal to century time scales (Coll et al., 2020; Lotze et al., 2019; Pörtner et al., 2022). Although this long term time scale is valuable for strategic planning, the majority of immediate political and commercial decisions are made on shorter time scales, the seasonal-to-decadal scale (Figure 1; Meehl et al., 2009; Payne et al., 2022).

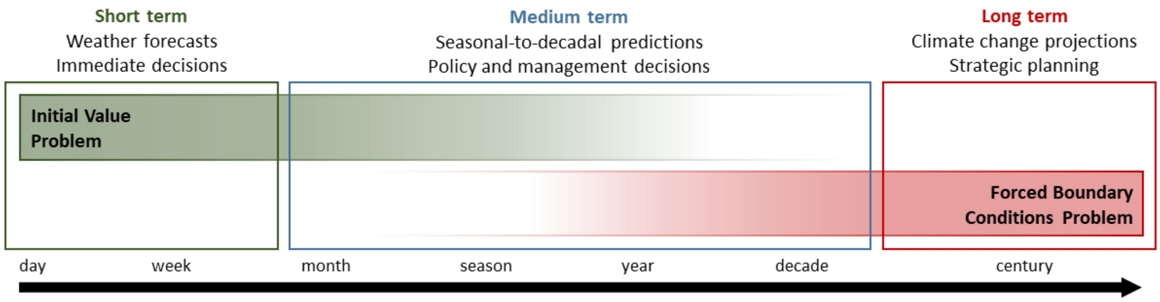

Figure 1 - Schematic of model- and decision time horizons, from short term forecasts to medium term predictions to long term projections. Short term forecasts are entirely dependent on starting conditions (the "initial value" problem), where long term projections are mostly affected by external drivers (the "forced boundary conditions" problem). Medium term predictions are affected by both problems (Adapted from Meehl et al. 2009).

At short time scales, from days up to a month, the predictive capacity of ocean and atmosphere models is firmly limited by the chaotic nature of the Earth system. Infinitesimal perturbations applied to a given set of initial conditions (the "initial value" problem, Collins, 2002; Meehl et al., 2009) lead to diverging trajectories in rather short temporal windows. On the other hand, at long time scales from decades to centuries, slow changes in external radiative forcings such as solar irradiance, aerosols and greenhouse gases (Meehl et al., 2009, the "boundary conditions problem", 2021) induce long-term trends that emerge over the chaotic variability. Since the pioneering studies of Smith et al. (2007), Keenlyside et al. (2008), and Pohlmann et al. (2009), the climate modelling community has been largely investing in improving the predictability on intermediate time scales, from months up to a decade, where climate models are both sensitive to initial value constraints and boundary conditions (Figure 1). This exercise has been underpinned by multi-model coordinated initiatives like the Decadal Climate Prediction Project (DCCP; Boer et al., 2016) and has been recently replicated with more complex Earth System Models (e.g., Ilyina et al., 2021; Li et al., 2016; Sospedra-Alfonso et al., 2021) capable of simulating, among other things, atmospheric chemistry and ocean biogeochemistry. These predictions rely on the initialization of the models with conditions that describe the best knowledge of a given observed state, a process that allows leveraging the predictability that arises from slow-paced internal variability processes, and are additionally driven with the historical and projected evolution of the main radiative forcing factors (e.g. solar irradiance, volcanic aerosols, concentrations of greenhouse gases) to capture the externally forced variability. The performance of these ESM-based predictions is evaluated by performing large sets of retrospective ensemble forecasts that are evaluated in terms of their ability to reproduce the observed variability. These predictions are typically initialised every year, and contain several ensemble members that are run forward for up to ten years (Figure 2; Boer et al., 2016).

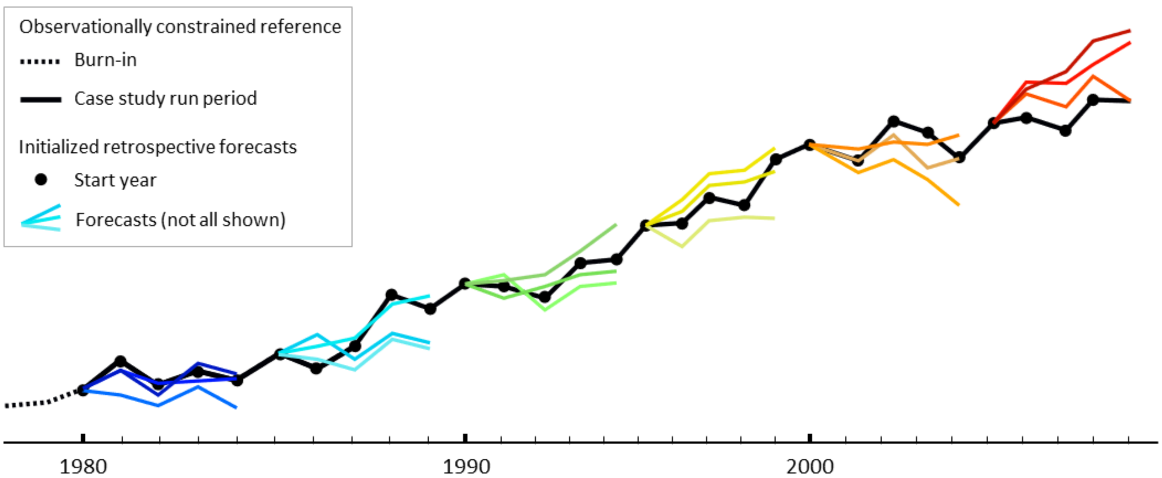

Figure 2 - Schematic of retrospective predictions to assess the impact of chaotic variability on the ability of ESMs to predict observations. The Y axis represents any dependant ESM variable included in retrospective predictions.

A next logical step is to assess whether/how predictive capacity of key ecosystem drivers within ESMs can significantly enhance the predictive skill in ecological models – a core scientific objective of EU Horizon 2020 project TRIATLAS (Tropical and South Atlantic Climate-Based Marine Ecosystem Prediction for Sustainable Management). To date, impacts of uncertainty related to the internal variability of ESMs on decadal time scales has been investigated for a handful of ecological hypotheses with encouraging results. For example, Årthun et al. (2018), Thorson (2019) and Payne et al. (2022) demonstrated improved confidence in predicting habitat suitability and species distribution shifts related to changes in ocean temperatures. Park et al. (2019) demonstrated that inter-annual variations in fish catches can be anticipated from ESM-based skilful predictions of phytoplankton and sea surface temperatures.

However, to our best knowledge, a systematic quantification of how ESM variability on decadal scales could cascade through complete marine food webs, and an evaluation of whether this variability has the potential to significantly change MEMs trajectories with the aim to improve the predictability of a MEM, have not yet been performed. Such an exercise would require systematically executing a MEM for potentially hundreds of retrospective forecasts, and analysing large volumes of spatial-temporal model output. This would require computing power far beyond a single workstation, and although the concept of using the combined power of a network of computers to solve demanding computational tasks dates at least back to the 1970's (e.g., Farber, 1970; Jones & Schwans, 1979; Vouk, 2008), the MEM community is mostly unable to utilise distributed computing power due to compounding challenges. Inherent limitations related to their computational complexity and structure, with long run times to represent non-linear processes at different temporal and spatial scales that cascade through food webs make MEMs incompatible with common high-performance computing technologies and computing scientific software execution infrastructures (Steenbeek et al., 2021). Scientific workflow management systems (Curcin et al., 2010; Wang et al., 2008), code execution frameworks (e.g., Ludescher et al., 2013), and commercial cloud computing solutions tend to require that hosted applications execute cleanly, safely, orderly and optimised by abiding to strict guidelines regarding programming languages and code architecture, execution efficiency, resource use, and scalability (e.g., Rimal et al., 2011). As MEMs are mostly developed on limited academic budgets with little involvement of IT staff, re-coding a MEM to match such requirements is too costly and perhaps even undesirable in order not to get locked into proprietary technological execution frameworks (Steenbeek et al., 2021). On the other hand, distributed computing via networked computers, virtual machines and virtualization technologies such as Kubernetes (Jeffery et al., 2021) and workload managers such as SLURM (Yoo et al., 2003) could certainly carry the systematic execution of ESM/MEM complexes in their original form, but require dedicated funding and technical support to operate and maintain. Whereas a few fortunate modellers may have access to institutional distributed computing environments and the dedicated staff to assist in the operation, the majority of the MEM community is left without practical solutions to systematically and comprehensively assess their models (Steenbeek et al., 2021).

The global MEM community needs a simple, generic and open-access framework that uses low-tech and free software to support the systematic mass-execution and mass-analysis of data- and computationally demanding scientific tools. Such a framework must allow the execution of software written in any language, as MEMs have been implemented in a broad range of platforms such as .NET, C, Fortran, Matlab, Python and R (e.g., Audzijonyte et al., 2019; Pal et al., 2020; Steenbeek et al., 2016). Such a framework must also support ecosystem modellers in deploying their workflows and toolkits in their original form. Ecosystem modelling is a complex field that combines understanding of marine biology and ecology, biochemistry, hydrology, fisheries dynamics and socio-economics, and that relies on the operation of a wide range of complex software tools to process, generate and analyse data. Thus, rather than requiring that analytical processes are translated into a common annotation, a scientific framework must acknowledge this diversity in software tools and support the execution of scientific workflows as they are. And last, to facilitate ease of use, the framework must seamlessly scale up desktop workflows across available hardware.

With these constraints met, such a framework would form the scaffolding for executing computationally demanding applications such as MEM validation, calibration and uncertainty assessments (Figure 3).

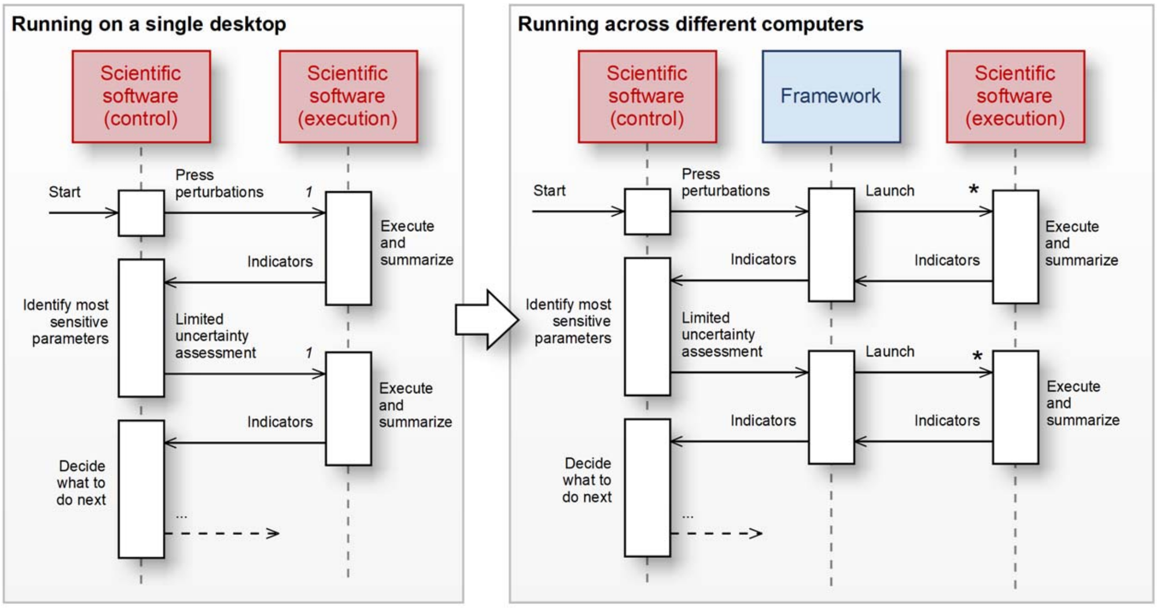

Figure 3 - A schematic overview of the workflow needed to systematically assess MEMs, here used to perform a hypothetical limited uncertainty assessment. The left panel shows this exercise deployed on a single desktop computer; the right panel shows this same exercise, transparently dispatched across any available hardware.

Here we present a prototype MEM run framework that we constructed to facilitate the systematic execution of marine ecosystem models. We apply this prototype framework to systematically run the seasonal-to-decadal retrospective predictions obtained from two different ESMs, EC-Earth3-CC and NorCPM1, through the mechistic, spatiotemporal explicit trophodynamic MEM EcoOcean. Through this, we demonstrate the feasibility of the approach as an important step toward making Marine Ecosystem Modelling operational.

2 Materials and Methods

Here we describe the main design considerations in developing the framework, and we present a case study to demonstrate that the framework can be used to systematically mass-execute MEMs. We then perform an indicative analysis to quantify whether ESM uncertainty has the potential to significantly affect the output of a complex and mechanistic global MEM, examining relevant functional groups within the food web in selected subregions of the global ocean. We outline the ESMs, the MEM and the runtime environment that we used, the application of the framework to perform the simulations, and a cursory analysis of modelling results.

2.1 Framework

The aim of the prototype MEM multi-run framework is to demonstrate that computationally heavy mechanistic and spatiotemporal MEMs can be systematically executed and analysed. Following the recommendations of Steenbeek et al. (2021), one should be able to operate the framework with minimal reliance on technical expertise, funding and specialised hardware to facilitate global uptake. Thus, instead of adopting an existing workflow management system, we opted to develop a framework from the ground-up to solely focus on the needed functionality without any additional complexity or restrictions related to funding and intellectual property.

The conceptual structure of the prototype MEM multi-run framework, henceforth referred to as “the framework”, is outlined in Figure 4.

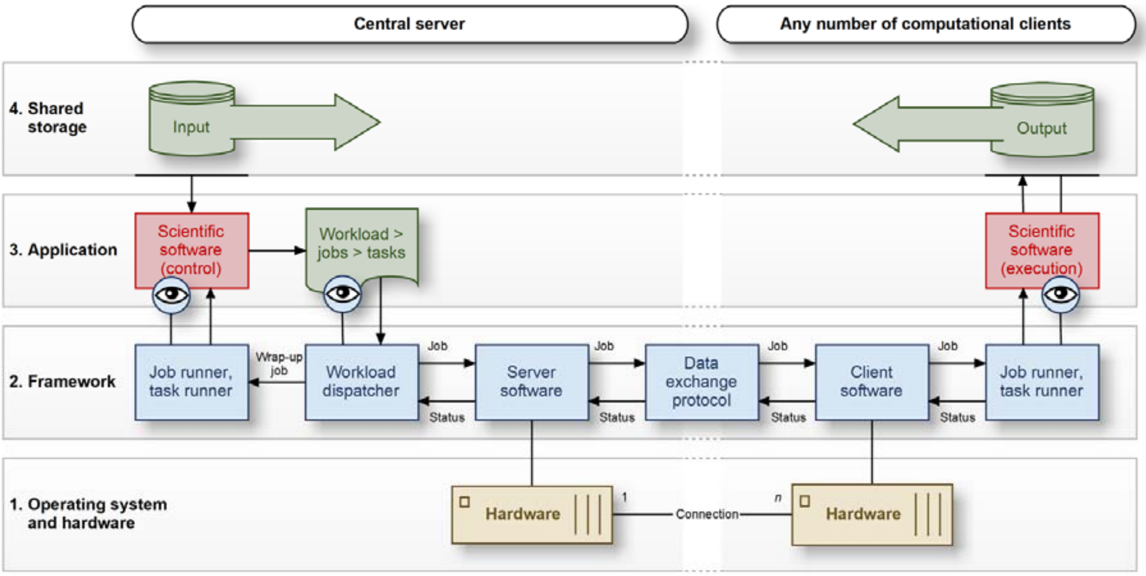

Figure 4 - The conceptual structure of the MEM multi-run framework. Set up as a server-client structure, the framework (2) dispatches the jobs that are defined within a scientific workload across available hardware (1). The framework loosely interacts with scientific software to execute the tasks within a job (3) and relies on available shared storage solutions (4) to distribute input data to clients, and collate resulting output on the server. The ‘eye’ icon reflects the loose interactions where the framework checks upon the state of external software and data without any form of technical integration and dependency. When a workload has been processed, scientific software is notified, which can dispatch a new scientific workload if desired.

Four independent and loosely connected layers (hardware, framework, application and shared storage) interplay as follows:

1. Hardware: The hardware layer can consist of any computing hardware able to run a particular MEM.

2. Framework: The framework layer handles the execution of scientific work across a computing network, and consists of the following components:

- A workload, which is a text file that describes the scientific work that the framework needs to execute. A workload consists of a number of independent computational experiments (Jobs), each in turn consisting of one or more executions of specific modelling scripts (Tasks). The workload also states which wrap-up job should be executed if the workload execution succeeds or fails. The wrap-up jobs provide the scientific application to decide on next execution steps such as dispatching a new workload. For a conceptual example of what a workload could look like, refer to Supplementary Material text S1, inset 1.

- A server, which is a small piece of software that maintains an active connection to available clients, with whom it can exchange information. Jobs in a workload are dispatched to clients;
- One or more clients, where the jobs are executed. Clients maintain an active connection to the central server and exchange information with it.
- Clients and the server can exchange information through a range of communication protocols built into the framework, each catering to different usage scenarios but that may require varying levels of IT expertise to deploy.
- A server-side work dispatcher handles and monitors workload execution: the dispatcher sends Jobs to clients, tracks their execution based on feedback from the clients, keeps track of the overall status of the workload execution, and upon completion, orders the server-side execution of the wrap-up job.
- A client-side job and task runner handles the sequential execution of the tasks within a job. Task execution involves starting scientific software, monitoring its progress, and waiting for its termination (or actively terminating it if scientific software has become unresponsive). Job and task execution status updates are sent back to the server. Security measures are in place to ensure that the framework only operates on pre-authorized folders and executables.

3. Application: The application layer consists of the scientific software that has been made available to the framework. Any software can be included as long as it can be parameterized and executed via a command line.

4. Shared storage: The shared storage layer makes sure that server-side input data is made available to client processes, and that scientific output generated at the client side is collated on the server. The prototype framework does not contain facilities to synchronise data, as there are plenty of viable solutions in the form of shared (network) storage, cloud storage providers, and file-sharing services.

Specific considerations

For ease of deployment and to demonstrate versatility, the prototype framework and any scientific application deployed across it are kept fully independent. Server-side scientific applications place workload text files in a predestined location for the work dispatcher to find. At workload execution completion, server-side wrap-up jobs can be used to activate the scientific applications once again to analyse execution results, and to dispatch a follow-up workload if desired.

Running jobs and tasks: The framework was made to launch two types of scientific applications.

The first category comprises the execution of stand-alone executables whose runtime behaviour can be controlled through the command line and that may execute a programmed script. The framework launches the executable and monitors its progress while capturing standard output and error information (Ritchie, 1984) to aid troubleshooting, and process exit codes (Maleki, 2022) to know whether a task succeeded or failed. If a stand-alone executable becomes unresponsive it can be terminated after a specified time-out. When the stand-alone executable terminates, exit codes with a value of zero indicate that the execution succeeded.

The second category includes internal executions - code that resides within the execution client - via a software engineering mechanism known as "runtime reflection" (Redondo et al., 2008; Schmidt et al., 2000, p. 134). Instead of indicating a physical separate executable, a task alias refers to the name of a specifically formed and recognizable piece of code that resides in the client code base and that implements a task. This code is dynamically looked up, executed and monitored for completion while standard output and error information (Ritchie, 1984) and the exit code are collected by the framework. In-client code is intended to facilitate running the framework on environments where clients cannot launch separate executables such as HPC clusters.

Information exchange: Because the framework can be deployed over operating systems (OS-es) that may be configured differently, scientific software deployed over the framework should use consistent natural language specific formatting of numbers, date- and time fields, and OS specifics such as text file line endings, etc.

Extensibility: In the particular case study outlined here, we implemented the framework in Microsoft Visual Basic.NET, compiled to .NET Standard 6.0 which produces executables that can be installed and natively executed on Windows and Linux OS-es. In order to customise the framework to future needs and to change and improve its functioning, the framework source code is organised as an open-source API that is open to modifications and extensions.

Installation and deployment: In order to use the framework, operators will need to prepare target computers with the framework software, cloud storage provider software, and the programs needed for the execution of a scientific workload. This is an unavoidable and possibly challenging task, but for use cases such as we present here where we interconnect regular desktop computers via cloud storage providers, this task should not be any more challenging than configuring a desktop computer for regular use.

For additional framework design considerations refer to Supplementary Material text S1.

2.2 Earth System Models

For this case study, two contrasting ESMs participating in TRIATLAS, EC-Earth3-CC and NorCPM, delivered estimates in photoplankton biomasses and sea water temperatures for the years 1950-2015. ESM variable names and units were standardized to the Climate Model Operator Rewriter (CMOR) standard 3.3 (Nadeau et al., 2018). Both models delivered a single continuous simulation reconstructing the evolution of the global biophysical system, and an ensemble of yearly starting retrospective predictions characterized by three arbitrarily selected members of their full ensemble. We used three members as a good trade-off to reasonably sample the ESM forecast uncertainty, while limiting the computational burden for the MEM simulations. Here we provide a brief technical summary of the two contrasting ESMs and their contribution to the case study. Considering two ESM models with different physical and biogeochemical ocean components, and for which the decadal predictions are also initialized in a different manner, offered the ability to explore the sensitivity of the MEM predictions to the uncertainties in the state variables used as boundary conditions.

EC-Earth3-CC (Döscher et al., 2022) is the ESM version of the global climate model EC-Earth that includes a description of the carbon cycle at its standard resolution. Its atmospheric component is the Integrated Forecast System (IFS) from the European Centre for Medium-Range Weather Forecasts (ECMWF) and uses a T255 horizontal resolution and 91 vertical levels. The ocean component is NEMO3.6 (Madec & the NEMO team, 2023), which includes the sea ice model LIM3 (Rousset et al., 2015) and the ocean biogeochemistry model PISCES (Aumont et al., 2015) integrated in the code. NEMO3.6 is run with an ORCA1 horizontal grid (i.e nominal one-degree horizontal resolution) and 75 vertical levels. Dynamical vegetation, land use, and terrestrial biogeochemistry are provided by LPJ-GUESS (B. Smith et al., 2014). The library OASIS3-MCT (A. Craig et al., 2017) is used for the coupling of most of the model's components. More detailed information on EC-Earth3-CC and its different components can be found in Döscher et al. (2022).

The predictions from EC-Earth3-CC were performed following the experimental protocol for the Decadal Climate Prediction Project (DCPP) experiments DCPP-A (Boer et al., 2016), with start dates for every 1st of November in the period 1980 to 2019. Start dates prior to 1980 were not included as the quality of the atmospheric/oceanic reanalysis used to initialise the ESM model cannot be properly validated for the pre-satellite era (i.e., before 1980) due to the lack of widespread biogeochemical observations. A total of 15 members were produced for each start date, with a forecast length of 7 years, instead of 10, to save computational resources.

The initialization protocol is a precursor of the methodology applied for the climate predictions of Bilbao et al. (2021). The ocean physical and biogeochemical conditions come from a reconstruction performed with the ocean component of EC-Earth3-CC (hereafter referred to as RECON) forced at the surface using an atmospheric reanalysis. In this reconstruction, observations for temperature and salinity are assimilated at the surface by adding fluxes for heat and freshwater to the energy and salinity conservation equations. At the same time, the interior of the ocean is also nudged towards a reference re-analysis product for both temperature and salinity. It is important to notice that no observations of ocean biogeochemistry or sea-ice are assimilated such that these fields are left free to evolve in response to ocean physics. More details about the EC-Earth3-CC initialization procedure as well as about the reference observation products used can be found in Supplementary Material text S2.

For this application, EC-Earth3-CC delivered monthly vertically integrated large and small phytoplankton carbon concentrations (lphyc and sphyc), and mean potential sea water temperatures (thetao) for the top 150m, the entire water column, and the bottom. These variables were delivered for a 1980-2015 continuous historical run (RECON) and for an ensemble of three 7-year retrospective predictions (i.e. r6i1p1f1, r7i1p1f1 and r8i1p1f1 DCPP-A members) with yearly start dates for the whole period 1980-2013.

NorCMP1, short for the Norwegian Climate Prediction Model version 1 (Bethke et al., 2021), is based on the Norwegian Earth System Model version 1 (NorESM1; Bentsen et al., 2013; Iversen et al., 2013) which is in turn based on the Community Climate System Model version 4 (CCSN4; Gent et al., 2010; Vertenstein et al., 2010) after important modifications. Its ocean component uses a standard horizontal grid (gx1v6) with 53 layers in an isopycnic vertical coordinate, which includes prognostic biogeochemical cycling in the form of the HAMburg Ocean Carbon Cycle (HAMOCC; Maier-Reimer, 1993; Maier-Reimer et al., 2005) adapted to this isopycnic ocean model framework (Tjiputra et al., 2010). The atmospheric component consists of the Oslo version of the Community Atmosphere Model (CAM4-OSLO; Kirkevåg et al., 2013), that has specialised chemistry-aerosol-cloud-radiation interaction schemes, with a two degree horizontal resolution and 26 levels in the vertical with a hybrid sigma-pressure coordinate. The land (same grid as the atmospheric component) and the sea ice (same grid as the ocean component) components are basically the same as in CCSM4, except for a scheme for dust deposition on snow/sea ice. The overarching execution control of the coupled system and the exchange of information between model components is handled by the CCSM4 coupler CPL7 (A. P. Craig et al., 2012). Detailed descriptions of the NorESM components and its biogeochemical ocean module can be found in Bentsen et al. (2013) and Tjiputra et al. (2010), respectively.

NorCPM1's DCPP-A simulations have start dates for every 15th of October in the period 1960 to 2018. A total of 10 members were produced for each start date, with a forecast length of 10 years (Bethke et al., 2021). Each member of these hindcast experiments ("hindcast-i2") are initialised by the 15 October states of the first 10 members of a data assimilation (DA) simulation ("assim-i2"), which uses oceanic observations to update ocean and sea ice components. This DA simulation uses a 1950-2010 SST reference climatology for computing anomalies, replacing the climatology of the observations by the model climatology calculated from the NorCPM1's 30-member no-assimilation historical experiment, and additionally updates the sea ice state via strongly coupled DA of the observations (Bethke et al., 2021). The DA scheme updates all ocean physical state variables but not the biogeochemical state variables. However, Fransner et al. (2020) showed that the initialization has no important effect on the predictability of ocean biogeochemistry beyond lead year 1, but also showed that assimilating SST can potentially constrain the near-surface primary production and hence the biogeochemical variability.

For this application, NorCPM1 delivered monthly mean integrated phytoplankton carbon (phyc), and mean potential sea water temperatures (thetao) for the top 150m, the entire water column, and the bottom. These variables were delivered for a 1980-2015 continuous historical run (HIST) and for an ensemble of three10-year retrospective predictions initialized every year in 1980-2008, which corresponds to members r1i2p1f1, r5i2p1f1 and r10i2p1f1 of the DCPP-A ensemble.

2.3 Marine Ecosystem Model

The MEM deployed in this case study is EcoOcean, a mechanistic, spatiotemporal ecosystem modelling complex of the global ocean that includes food-web dynamics from primary producers to top predators under influence of anthropogenic activities and climate change. EcoOcean has at its core the Ecopath with Ecosim (EwE) modelling approach (Christensen & Walters, 2004), where the spatial-temporal module Ecospace has been heavily modified to represent spatial heterogenity in fishing and the behaviour, growth and movement of functional groups across the worlds' oceans (Christensen et al., 2015; Coll et al., 2020).

EcoOcean was parameterized and calibrated as described in Coll et al. (2020), as used for the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP) simulation round 2b to explore how projected climate change might affect future (2016-2100) ocean ecosystems (Tittensor et al., 2021). The EcoOcean MEM operates on a spatial grid of one decimal degree at monthly time steps with a food web that consists of 52 interconnected functional groups. Functional groups are represented spatially accounting for approximately 3400 species that underpin the functional groups. Functional groups disperse, gravitating towards cells with more suitable feeding conditions and lower risks of depredation, where feeding suitability is determined by the Ecospace habitat foraging capacity model (Christensen et al., 2014) modified by cell-specific responses, temperature-adjusted metabolic rates, and species' native ranges to constrain the initial distribution of functional groups to observed occurrences (Coll et al., 2020). Fishing is driven by historical effort (1950-2015) for 14 fleets (Rousseau et al., 2019). Historical fishing effort is introduced as a total per each of the 66 Large Marine Ecosystems, within each fishing effort is distributed via a simple gravity model that considers the distributions and market value of targeted functional groups versus the cost of fishing in any given location that is not closed to fishing (Christensen et al., 2015).

Relevant to this case study is how EcoOcean utilises the Earth System model output to drive its global ecosystem dynamics. EcoOcean contains three functional groups of phytoplankton: large, small and diazotrophs, that alongside benthic producers and bacteria act as the nutritional foundation for the food web. When connected to global Earth System Models, EcoOcean typically overwrites its spatially distributed phytoplankton biomasses with ESM-delivered phytoplankton biomass for matching timesteps (Coll et al., 2020; Steenbeek et al., 2013; Tittensor et al., 2018), scaled to 1950 biomass estimates EcoOcean was calibrated to. Furthermore, EcoOcean v2 (Coll et al., 2020) linked sea surface temperature to affect functional group productivity and distributions via the built-in habitat foraging capacity model (Christensen et al., 2014). For this case study, functional group responsiveness to climate was extended by associating pelagic, benthopelagic and demersal functional groups with mean temperatures for the top 150m, entire water column and bottom, respectively. This refinement was made to capture temperature fluctuations at depth as delivered by the ESM retrospective predictions.

2.4 Runtime environment

The framework was deployed across a network of computers with varying specifications as shown in Table 1. All computers hosted 64-bit Operating Systems and were powerful enough to execute EcoOcean. Machines were located in two physical locations, interlinked via a Dropbox (www.dropbox.com) professional plan with 2TB of storage space for mass data transfer, and a free Sync (www.sync.com) account for framework communication. For every 4 threads or fewer, a separate framework client was created, which meant that the runtime environment was able to simultaneously perform 24 executions of EcoOcean (Table 1, Nº clients).

Table 1 - the computers used to perform the case study, with key characteristics

Computer	Year	OS	Processors	Nº threads	Nº clients
Desktop	2013	Win 7	2 x i5 quad-core	8	2
Laptop	2015	Win 10	1 x i7 quad-core	4	1
Laptop	2018	Ubuntu	2 x i7 quad-core	8	2
Desktop	2020	Win 10	2 x i7 quad-core	8	2
Tower	2022	Win 10	20 x i7 quad-core	48	12
Laptop	2023	Win 11	14 x i7 quad-core	20	5
Total					24

2.5 Application

The EcoOcean executions were encapsulated in a custom developed command-line utility, henceforth referred to as the "EcoOcean wrapper", that configured the EcoOcean model for executing a specific simulation, executed the simulation, intercepted and condensed EcoOcean maps over time into time series, and saved these time series into one ZIP file per run. By specifying a ZIP output file name adhering to a simple and strict naming protocol, the command line utility understood exactly how to configure and run EcoOcean, how to name the output ZIP file, and how to finally place these output ZIP files directly into a Dropbox folder dedicated to multi-run framework server-side data collation.

All EcoOcean simulations started in the year 1950 after a ten-year spin-up (or burn-in) period, and were executed through 2015. EcoOcean output was collated for the period 1980-2015. For both ESMs, EcoOcean was executed with and without fishing following Coll et al. (2020). ESM data were delivered to EcoOcean in the form of monthly varying maps. The maps representing mean temperatures for the top 150m, water column and bottom were fed into the EcoOcean habitat foraging capacity model (Christensen et al., 2014), and the maps for large and small phytoplankton were used to force the magnitudes and distributions of the corresponding phytoplankton groups within EcoOcean (Coll et al., 2020).

For the two ESMs and the two fishing scenarios, EcoOcean was driven by ESM historical data to gather simulation baseline output. Then, for the two ESMs, two fishing scenarios and every retrospective prediction start year for the three members, EcoOcean was executed with historical data up to the start year of a retrospective prediction, after which EcoOcean was executed until the end of the retrospective predictions while being driven by the ESM data for that retrospective prediction. For the retrospectice prediction experiments, output was only collected for the period covered by the 7- (EC-Earth3-CC) or 10 (NorCPM1) year retrospective predictions.

As EC-Earth3-CC data started at 1980, historical data for the year 1980 were repeated during the EcoOcean spin-up period and for the period from 1950 through 1980. NorCPM1 did not distinguish explicitly between small and large phytoplankton; therefore, the total phytoplankton biomass data was used to proportionally drive large and small phytoplankton dynamics in EcoOcean.

2.6 Analysis

EcoOcean produced global 1 degree gridded maps of biomass and catch (where applicable) by functional group at monthly time steps, which can produce a file volume upward of 50GB per simulation. To save storage space while retaining important signals, we condensed EcoOcean output into time series for the hydrological basins of the world (Figure 5; FAO, 2020) and the major fishing areas for statistical purposes (Figure 6; FAO, 2015) as defined by the Food and Agricultural Organization (FAO). Each time series described the mean biomass and catch, per functional group and per region, weighted by cell area. The use of regional time series was decided on as an effort to capture regional variability in ecosystem dynamics for MEM run comparison whilst significantly reducing the volume of model output transferred and analysed.

2 North Atlantic

3 Central Atlantic, Mediterranean and the Black Sea

4 South Atlantic

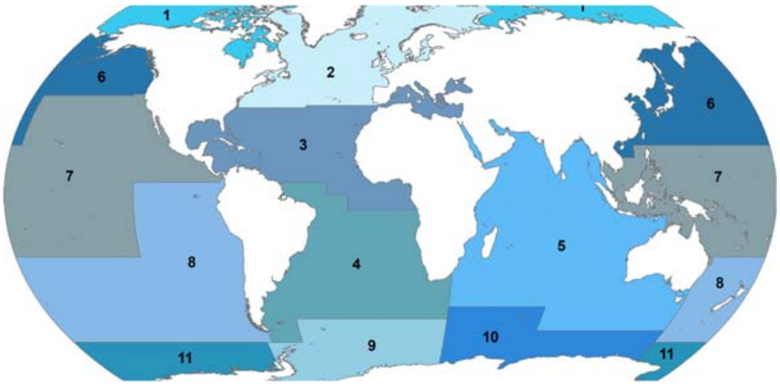

Figure 5 - The eleven ocean sub-basins as defined by FAO, used in the 384 MEM simulations to summarise trends in functional group catches and biomasses. In this manuscript, only areas 2, 3 and 4 (north-, central- and south Atlantic) are presented.

21 NW Atlantic

27 NE Atlantic

31 CW Atlantic

34 CE Atlantic

41 SW Atlantic

47 SE Atlantic

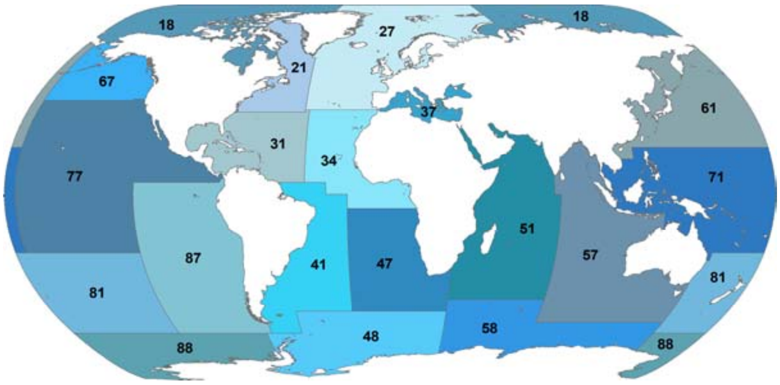

Figure 6 - The fishing areas for statistical purposes as defined by FAO, used in the 384 MEM simulations to summarise trends in functional group catches and biomasses. In this manuscript, only areas 21 and 27 (northwest- and northeast-), 31 and 34 (centralwest- and centraleast-) and 41 and 47 (southwest- and southeast Atlantic) are presented.

Although EcoOcean produced global results for 51 functional groups, this prototype case study focused on trends in biomass for only 6 functional groups: small, medium and large pelagic fish, and small, medium and large demersal fish. The choice of small, medium and large fish would allow for detecting direct changes induced by phytoplankton variability (small fish) and trophic cascades (medium and large fish). Different vertical positioning of selected functional groups could reveal relevant effects at depth. All comparisons were made for fished and non-fished MEM executions.

Results were analysed for three FAO sub basins: the north, central and south Atlantic, in line with the aims of EU Horizon 2020 project TRIATLAS (Figure 5).

2.7 Statistical measures

We explored the utility of a number of simple statistical measures to quantify how ESM uncertainty affects the output of EcoOcean when compared to output generated via the ESM baseline runs. In the formulae below, *n* = number of observations y_i; y_i = *observations* (EcoOcean output driven by the ESM baselines); $\hat{y}_i$ = *estimations* (EcoOcean output-driven by ESM retrospective predictions).

All statistical measures were calculated for three pelagic and three demersal fish functional groups, for both ESMs, for the three TRIATLAS regions (North, Central and South Atlantic), under fished and non-fished scenarios.

1. Root Mean Squared Error or RMSE (equation 1) measures the average magnitude (t/km^2) of the differences between predicted values and observed values. A lower RMSE indicates better predictive performance. It penalises larger errors more heavily than smaller errors due to the squaring of the errors. RMSE is sensitive to outliers since it squares the errors. RMSE is therefore a useful metric to quantify for which ecosystem components, and in which regions, ESM uncertainty mostly affects the marine ecosystem. Such outliers could indicate direct sensitivities to small perturbations, or could indicate ecosystem cascades.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad \text{Eq. 1}$$

2. Mean Absolute Error or MAE (equation 2) measures the average magnitude (t/km^2) of the absolute differences between predicted values and observed values. Like RMSE, a lower MAE indicates better predictive performance. MAE treats all errors equally and is not as sensitive to outliers as RMSE. MAE is a useful metric to quantify where ESM uncertainty has less impact on MEM predictions.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad \text{Eq. 2}$$

3. Symmetric Mean Absolute Percentage Error or SMAPE (equation 3) measures the percentage difference between predicted values and observed values, averaged across all observations. It is symmetric because it considers both overestimations and underestimations equally. SMAPE is easy to interpret in percentage terms and is suitable when dealing with data with varying scales. Because SMAPE ignores scale and direction, it is a useful metric to directly compare the relative error, directly or indirectly caused by ESM uncertainty, between functional group predictions for the historical runs and for the runs executed with ESM uncertainty for all regions.

$$SMAPE = \left(\frac{1}{n}\sum_{i-1}^{n}\frac{|y_i - \hat{y}_i|}{(|y_i| + |\hat{y}_i|)}\right) \times 100 \qquad \text{Eq. 3}$$

4. Pearson's Correlation Coefficient (equation 4) measures the linear relationship between predicted values and observed values. It ranges from -1 to 1, where 1 indicates a perfect positive linear correlation, -1 indicates a perfect negative linear correlation, and 0 indicates no linear correlation. A higher absolute value of the correlation coefficient suggests a stronger linear relationship between predictions and observations. Additionally, the Pearson coefficient can reveal hidden correlations for data that are not normally distributed. This coefficient is thus useful in correlating the linearity between historically- and uncertainty-driven MEM simulations, indicating where significant deviations may require further study.

$$r = \frac{\sum_{i=1}^{n}(y_i - \bar{y})\left(\hat{y}_i - \bar{\hat{y}}\right)}{\sum_{i=1}^{n}(y_i - \bar{y})^2\sum_{i=1}^{n}\left(\hat{y}_i - \bar{\hat{y}}\right)^2} \qquad \text{Eq. 4}$$

5. Directional Symmetry or DS (equation 5) measures the percentage of occurrences where the sign, positive or negative, of an observed and a predicted time series is the same. This coefficient is useful to correlate the direction of change between historically- and uncertainty-driven MEM simulations.

$$DS(y,\hat{y}) = \frac{100}{n-1}\sum_{i=1}^{n} d_i,$$
$$d_i = \begin{Bmatrix} 1, & if\ (y_i - y_{i-1})(\hat{y}_i - \hat{y}_{i-1}) > 0 \\ 0, & otherwise \end{Bmatrix} \qquad \text{Eq. 5}$$

4 Results

4.1 Framework performance

The prediction experiments resulted in a workload of 384 jobs, each job containing only one task: the invocation of the EcoOcean wrapper command-line utility. The driver data delivered by both ESMs comprised approximately 1 million time-tagged maps at a volume just over 415 GB. EcoOcean produced an estimated volume of 5 TB in output maps that were condensed into time series CSV files by the EcoOcean execution wrapper on the framework clients. The EcoOcean wrapper then compressed the time series CSV files and placed them in the Dropbox output folder for automatic transport to the framework server computer. By using time series, the framework produced a more manageable output volume of 50 GB, which was compressed to 3GB for file transfer to the server for analysis. The full set of EcoOcean simulations required approximately 2600 hours of CPU time, but via the framework used here - with a total of 164 computational cores (Table 1) - the complete set of simulations was performed in just under 30 hours.

The stability of the framework was assessed by randomly stopping and starting, and randomly adding and removing, computational clients during extensive test runs. The framework recovered from the resulting communication failures within a few minutes, rescheduling interrupted model executions or dispatching work to newly available clients. The use of cloud storage providers for main communication transport was slow but quite reliable. On a few rare occasions, the cloud storage providers stopped synchronising information entirely, which is an acknowledged remote possibility for both Dropbox (Dropbox.com, 2023) and Sync (Sync.com, 2023). In such cases, the affected client computers were no longer able to participate in a particular simulation run until their local cloud daemons were manually restarted. In one particular simulation test run, the framework server daemon stopped synchronising, which effectively terminated the entire experiment since the framework does not (yet) feature server redundancy.

The 384 ecosystem model executions functioned as expected, without errors in accessing and integrating ESM data into the running model, executing the model, extracting and collating output, and placing the output in the desired, pre-configured output locations.

4.2 Simulations

The simulations provided four sets of output - for the two different ESMs under fished and non-fished oceans, each featuring time series trends for the 11 ocean sub-basins and 19 ocean statistical areas for fisheries purposes, for 52 functional groups. Figure 7 shows what these data look like when plotted.

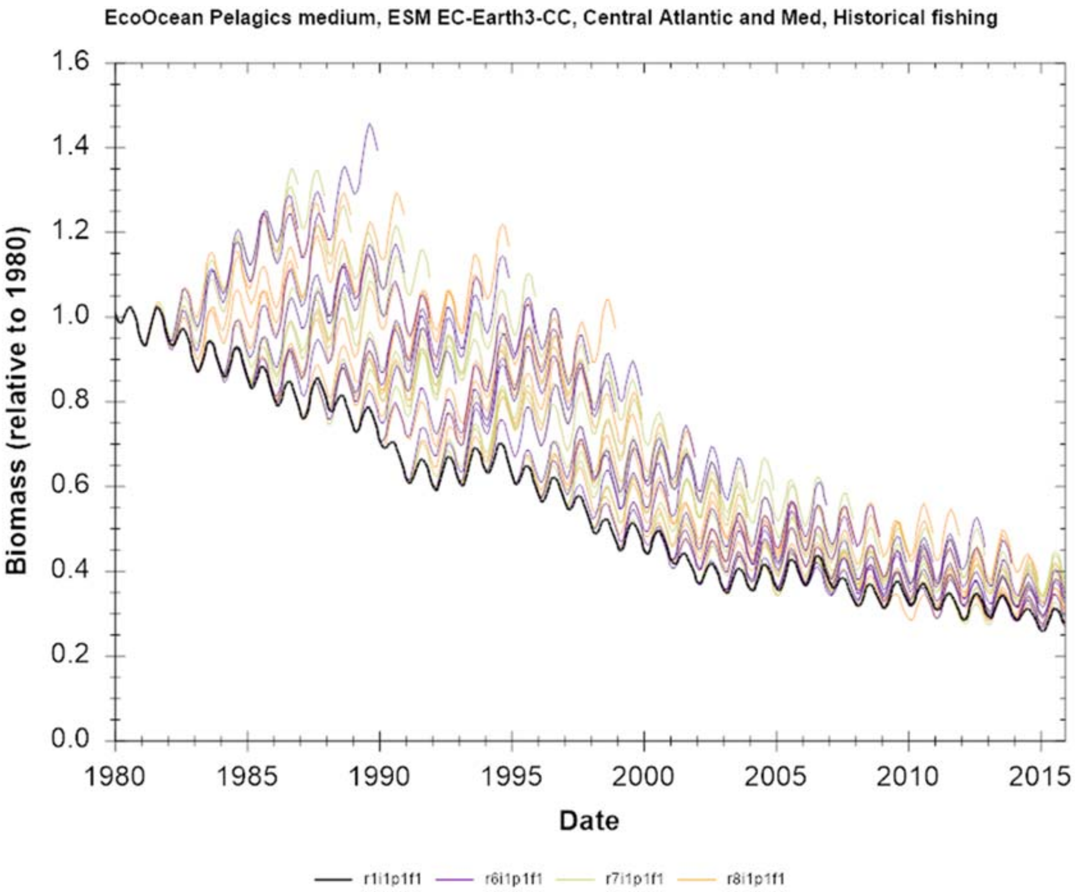

Figure 7 - An example of EcoOcean estimates for medium pelagic fish in the central Atlantic and Mediterranean, when the MEM is driven by output from EC-Earth3-CC under historical fishing pressure. The black line represents the EcoOcean output when driven by the continuous

ESM baseline, and the three coloured lines represent EcoOcean estimates when deviating away from the baseline for 7-year retrospective predictions. Ecosystem output is plotted relative to the annual average 1980 value.

Overall, results show that across the food web and the observed regions, the EcoOcean biomass trajectories displayed varying degrees of responsiveness to ESM uncertainty, depending on position of the selected functional group in the EcoOcean food web, the presence of fishing, the region analysed, and choice of earth system model linked to EcoOcean. For instance, Figure 8 and Figure 9 show EcoOcean estimates when driven by EC-Earth3-CC and NorCPM1 respectively, for the same functional group, large pelagic fish, which encompasses dolphinfish, sailfish, tuna, mackerel, marlin, swordfish and others. From these plots, a few things become clear.

- The impact of retrospective predictions for EC-Earth3-CC (Figure 8) tends to deviate from the observationally-constrained reference, while the impact of retrospective predictions for NorCPM1 (Figure 9) centres around the baseline r1i1p1f1 simulation.
- Fishing severely impacts large pelagic fish, regardless of ESM selected.
- Although the overarching trends are similar between the two ESMs, fishing has a much stronger relative impact on large pelagic fish in the Central Atlantic when EcoOcean is driven with EC-Earth3-CC output than with NorCPM1.
- NorCPM1 appears to introduce higher seasonal variability than EC-Earth3-CC, but this is probably an artifact of driving both small and large phytoplankton with the same single NorCPM1 phytoplankton estimates. Having both large and small phytoplankton follow exactly the same trend is expected to exaggerate the impact of phytoplankton fluctuations onto the EcoOcean food web, which in the case of EC-Earth does not happen as large and small plankton compete for the same nutrients.

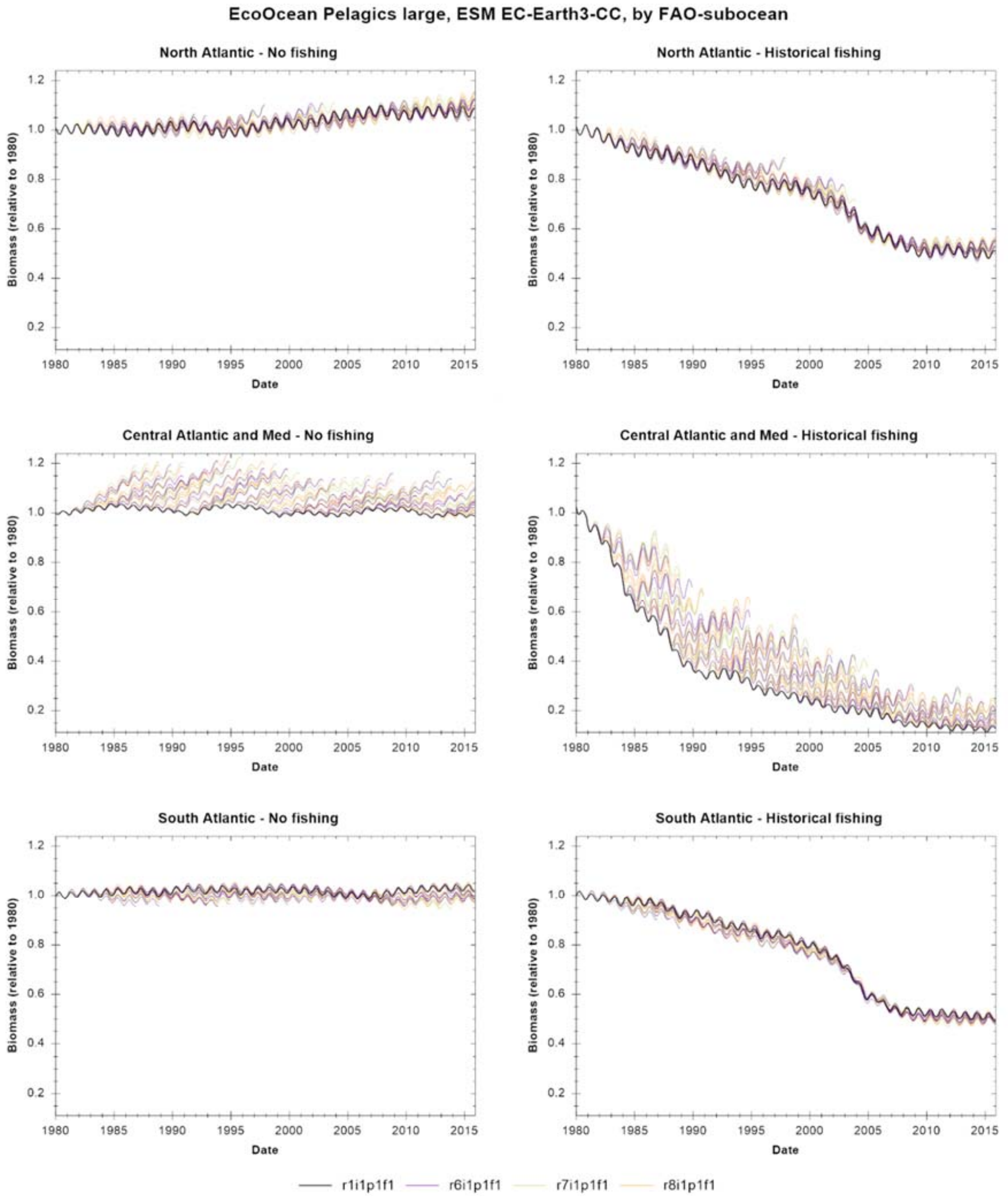

Figure 8 - Average EcoOcean biomass trends for large pelagic fish in three selected Atlantic FAO sub-ocean regions, for the years 1980-2015, when the MEM is driven by EC-Earth-CC historical data r1i1p1f1 (black line, one continuous EcoOcean run) and realisations r6i1p1f1, r7i1p1f1, and r8i1p1f1 (coloured lines). The left column shows EcoOcean biomass trends

without fishing, the right column includes historical fishing. All plots scale relative to their 1980 annual mean to standardise axes and to highlight the relative trends.

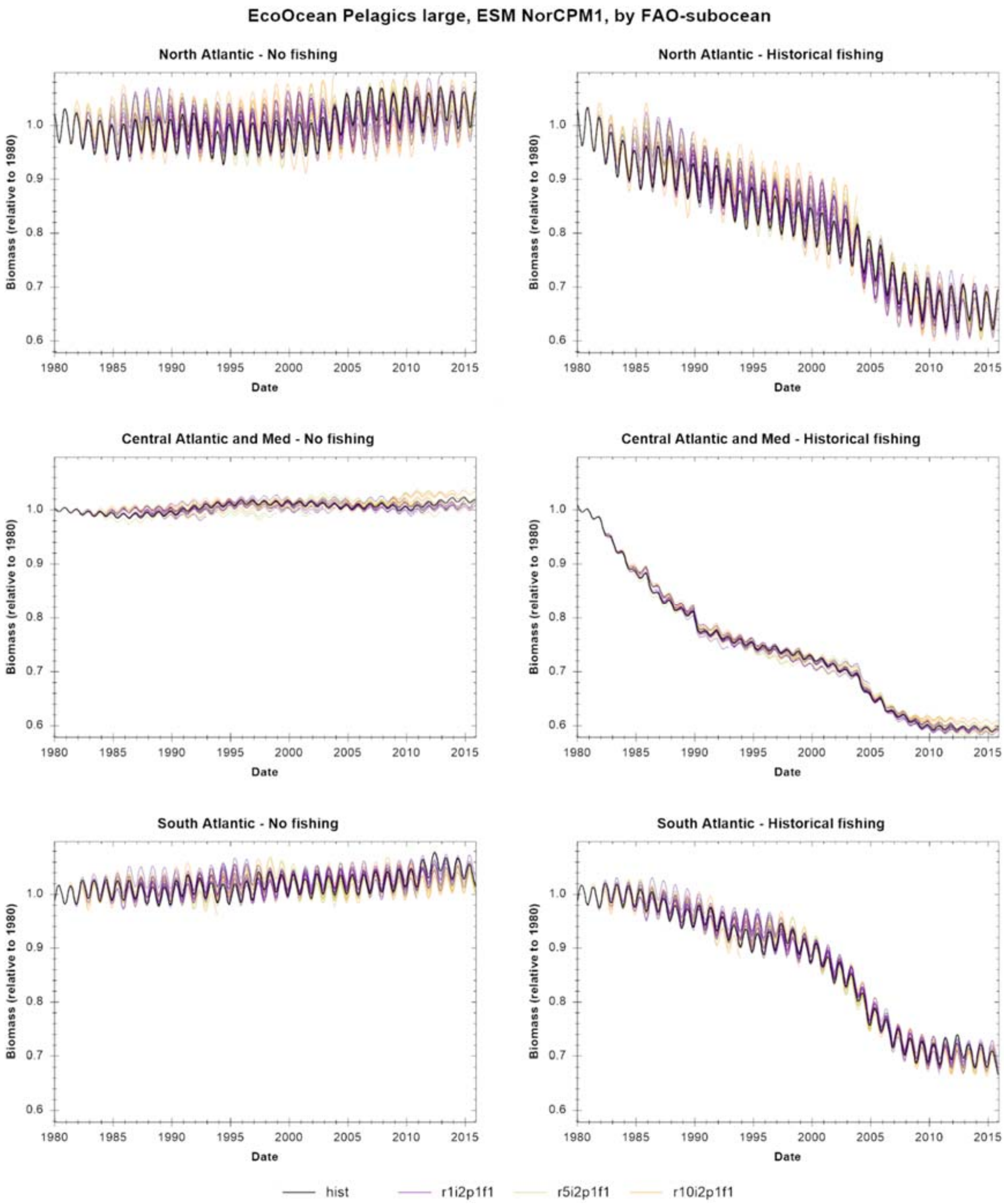

Figure 9 - Average EcoOcean biomass trends for large pelagic fish in three selected Atlantic FAO sub-ocean regions, for the years 1980-2015, when the MEM is driven by NorCPM1 historical data (black line, one continuous EcoOcean run) and realisations r1i2p1f1, r5i2p1f1, and r10i2p1f1 (coloured lines). The left column shows EcoOcean biomass trends without fishing,

the right column includes historical fishing. All plots scale relative to their 1980 annual mean, with standardised scales to highlight the relative trends.

The statistical measures (Table 2 and Table 3) captured these differences, comparing the last five years of the baseline simulation ("observations") against the mean model output for the retrospective predictions ("predictions"):

- For both Pelagic and Demersal components, the Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) were lower for EC-Earth3-CC than NorCPM1, indicating that in absolute terms, EcoOcean output was less affected by internal sensitivity of EC-Earth3-CC than NorCPM1;
- On the other hand, the Symmetric Mean Absolute Percentage Error (SMAPE) was generally lower for NorCPM1 than for the simulations driven by EC-Earth-CC, indicating that the trends produced by EcoOcean were less sensitive to internal uncertainty in the scenarios driven by NorCPM1 than EC-Earth3-CC;
- The Pearson's Correlation Coefficient was higher under fishing scenarios, regardless of ESM used. This indicates that fishing has a much stronger impact on EcoOcean output than ESM internal variability;
- Last, Directional Symmetry was higher for NorCPM1 than for EC-Earth3-CC, indicating that observations and predictions were generally more directionally aligned for NorCPM1 than EC-Earth3-CC.

Please note that RMSE and MAE measure absolute errores while SMAPE measures relative errors, which is reflected in Tables 1 and 2 in the range differences in all the categories.

Efforts to relate changes in ESM drivers to the various MEM outputs did not yield any useful signals, and will require a systematic attribution investigation.

Table 2 - Statistical measures to capture pelagic fish temporal biomass dynamics for 2010-2015.

PELAGIC BIOMASS

Earth.System	Functional.Group	Area.category	rmse	mae	smape	pearson	ds
Fishing							
ec-earth	Pelagics small	North Atlantic	0.00003	0.00003	0.306	0.95	93.8
ec-earth	Pelagics small	Central Atlantic, Med	0.00002	0.00002	11.406	0.91	90.7
ec-earth	Pelagics small	South Atlantic	0.00003	0.00003	2.107	0.90	98.8
ec-earth	Pelagics medium	North Atlantic	0.00016	0.00014	0.400	1.00	87.6
ec-earth	Pelagics medium	Central Atlantic, Med	0.00005	0.00004	9.805	0.98	93.6
ec-earth	Pelagics medium	South Atlantic	0.00010	0.00010	1.457	1.00	96.7
ec-earth	Pelagics large	North Atlantic	0.00001	0.00001	0.785	1.00	96.2
ec-earth	Pelagics large	Central Atlantic, Med	0.00000	0.00000	12.188	1.00	91.7
ec-earth	Pelagics large	South Atlantic	0.00000	0.00000	0.665	1.00	94.5
norcpm1	Pelagics small	North Atlantic	2.18926	1.83771	1.699	0.97	100.0
norcpm1	Pelagics small	Central Atlantic, Med	0.32014	0.24553	0.546	0.94	98.8
norcpm1	Pelagics small	South Atlantic	0.32270	0.26224	0.637	0.96	96.0
norcpm1	Pelagics medium	North Atlantic	2.72220	2.15267	0.998	0.95	98.3
norcpm1	Pelagics medium	Central Atlantic, Med	0.27593	0.20860	0.460	1.00	98.6
norcpm1	Pelagics medium	South Atlantic	0.52388	0.42735	0.597	0.96	97.4
norcpm1	Pelagics large	North Atlantic	0.07741	0.06563	0.793	1.00	96.0
norcpm1	Pelagics large	Central Atlantic, Med	0.01214	0.00981	0.177	1.00	96.9
norcpm1	Pelagics large	South Atlantic	0.04663	0.03321	0.335	0.99	97.4
No fishing							
ec-earth	Pelagics small	North Atlantic	0.00003	0.00003	0.270	0.99	89.5
ec-earth	Pelagics small	Central Atlantic, Med	0.00003	0.00003	7.971	0.71	88.6
ec-earth	Pelagics small	South Atlantic	0.00003	0.00003	2.413	0.80	98.6
ec-earth	Pelagics medium	North Atlantic	0.00019	0.00017	0.230	1.00	88.1
ec-earth	Pelagics medium	Central Atlantic, Med	0.00011	0.00010	3.760	0.79	86.0
ec-earth	Pelagics medium	South Atlantic	0.00018	0.00018	1.848	0.98	89.5
ec-earth	Pelagics large	North Atlantic	0.00001	0.00001	0.410	0.97	96.4
ec-earth	Pelagics large	Central Atlantic, Med	0.00001	0.00001	2.508	0.60	86.7
ec-earth	Pelagics large	South Atlantic	0.00001	0.00001	0.573	0.64	97.4
norcpm1	Pelagics small	North Atlantic	2.24326	1.88046	1.670	0.98	100.0
norcpm1	Pelagics small	Central Atlantic, Med	0.33540	0.25589	0.546	0.91	99.3
norcpm1	Pelagics small	South Atlantic	0.32477	0.26395	0.635	0.95	96.0
norcpm1	Pelagics medium	North Atlantic	3.08340	2.42889	0.900	0.94	99.1
norcpm1	Pelagics medium	Central Atlantic, Med	0.37913	0.26935	0.407	0.79	98.8
norcpm1	Pelagics medium	South Atlantic	0.55232	0.45123	0.582	0.84	96.7
norcpm1	Pelagics large	North Atlantic	0.09457	0.08470	0.712	0.89	99.1
norcpm1	Pelagics large	Central Atlantic, Med	0.01745	0.01409	0.141	0.91	93.1
norcpm1	Pelagics large	South Atlantic	0.05654	0.03972	0.318	0.91	98.3

Table 3 - Statistical measures to capture demersal fish temporal biomass dynamics for 2010-2015

DEMERSAL BIOMASS

Earth.System	Functional.Group	Area.category	rmse	mae	smape	pearson	ds
Fishing							
ec-earth	Demersals large	North Atlantic	0.00000	0.00000	5.55	0.94	92.4
ec-earth	Demersals large	Central Atlantic, Med	0.00008	0.00007	2.89	0.92	96.7
ec-earth	Demersals large	South Atlantic	0.00004	0.00003	2.25	1.00	88.1
ec-earth	Demersals medium	North Atlantic	0.00000	0.00000	9.28	1.00	84.8
ec-earth	Demersals medium	Central Atlantic, Med	0.00004	0.00003	1.97	0.99	77.9
ec-earth	Demersals medium	South Atlantic	0.00002	0.00001	1.80	1.00	88.4
ec-earth	Demersals small	North Atlantic	0.00000	0.00000	0.13	1.00	100.0
ec-earth	Demersals small	Central Atlantic, Med	0.00000	0.00000	5.86	1.00	85.5
ec-earth	Demersals small	South Atlantic	0.00000	0.00000	2.71	1.00	94.3
norcpm1	Demersals large	North Atlantic	0.69205	0.55441	1.18	0.98	100.0
norcpm1	Demersals large	Central Atlantic, Med	0.15360	0.11408	0.67	0.99	98.6
norcpm1	Demersals large	South Atlantic	0.18656	0.15419	0.99	0.96	98.8
norcpm1	Demersals medium	North Atlantic	0.30400	0.26852	2.56	0.76	95.7
norcpm1	Demersals medium	Central Atlantic, Med	0.01985	0.01497	0.54	0.99	94.8
norcpm1	Demersals medium	South Atlantic	0.02425	0.02010	0.76	0.90	96.4
norcpm1	Demersals small	North Atlantic	0.02187	0.01670	0.63	1.00	96.2
norcpm1	Demersals small	Central Atlantic, Med	0.00387	0.00282	0.49	1.00	96.2
norcpm1	Demersals small	South Atlantic	0.00652	0.00507	0.43	1.00	96.9
No fishing							
ec-earth	Demersals large	North Atlantic	0.00000	0.00000	2.14	0.96	91.4
ec-earth	Demersals large	Central Atlantic, Med	0.00010	0.00008	2.44	0.90	98.1
ec-earth	Demersals large	South Atlantic	0.00007	0.00006	1.89	0.63	79.3
ec-earth	Demersals medium	North Atlantic	0.00000	0.00000	5.52	0.75	85.3
ec-earth	Demersals medium	Central Atlantic, Med	0.00006	0.00005	1.36	0.94	77.0
ec-earth	Demersals medium	South Atlantic	0.00003	0.00003	1.79	0.78	85.7
ec-earth	Demersals small	North Atlantic	0.00000	0.00000	2.79	0.92	77.9
ec-earth	Demersals small	Central Atlantic, Med	0.00001	0.00001	0.49	0.98	82.9
ec-earth	Demersals small	South Atlantic	0.00001	0.00001	1.68	0.67	58.7
norcpm1	Demersals large	North Atlantic	0.83073	0.66392	1.16	0.98	99.8
norcpm1	Demersals large	Central Atlantic, Med	0.18570	0.13860	0.65	0.92	97.2
norcpm1	Demersals large	South Atlantic	0.19258	0.16033	0.97	0.92	98.6
norcpm1	Demersals medium	North Atlantic	0.32651	0.28874	2.40	0.82	94.5
norcpm1	Demersals medium	Central Atlantic, Med	0.02382	0.01851	0.53	0.82	94.5
norcpm1	Demersals medium	South Atlantic	0.02518	0.02084	0.74	0.59	96.7
norcpm1	Demersals small	North Atlantic	0.07640	0.05629	0.56	0.98	98.6
norcpm1	Demersals small	Central Atlantic, Med	0.01205	0.00928	0.22	0.74	95.3
norcpm1	Demersals small	South Atlantic	0.01113	0.00806	0.26	0.47	95.5

A side-by-side comparison of ecosystem trends for regions at different scale shows how different aggregation regions may reveal quite different trends (Figure 10). All side-by-side comparison plots (Figures S1-S36) are included in the supplementary material to indicate the vast spread of variation that emerges when aggregating MEM output over spatial areas with different sizes.

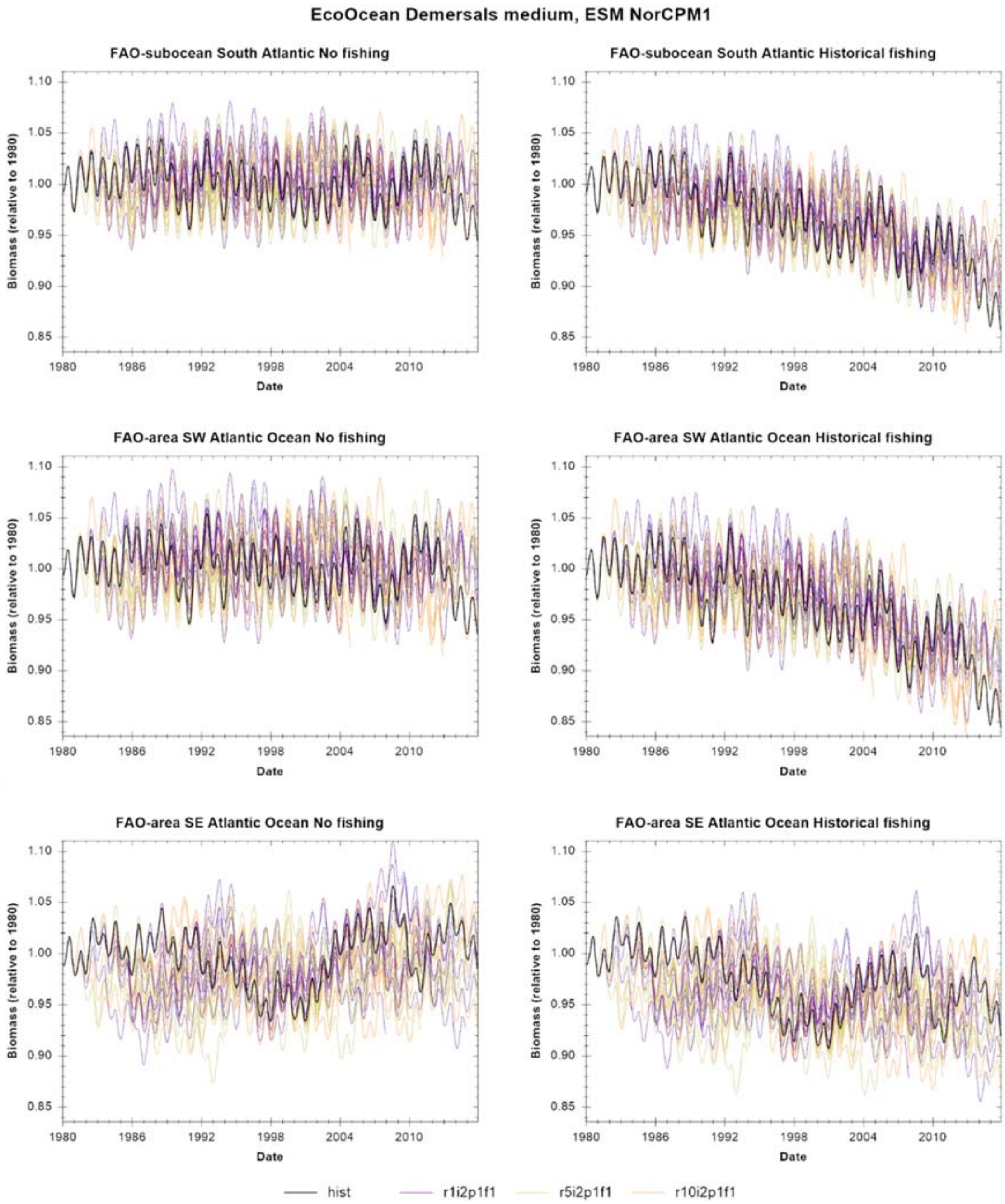

Figure 10 - Medium demersal fish biomass time series as predicted by EcoOcean when driven by NorCPM1. The plots show time series for FAO subocean south Atlantic (top row), and for the two subdivisions of that subocean, the southwest Atlantic (middle row) and the southeast

Atlantic (bottom row). Time series are shown without fisheries (left column) and with historical fisheries (right column).

5 Conclusions

In this study, we demonstrated that a distributed run framework built from simple technologies can be used to systematically run a marine ecosystem model, paving the way for systematic assessments that, prior, were deemed impossible to those without access to well supported powerful hardware and programming experience.

5.1 Experience using the framework

The framework performed well, despite its conceptual simplicity and reliance on the most basic technologies. The use of cloud storage providers for framework communication was not without caveats. We started out by using one storage provider, Dropbox, to handle all server-client data transfer, but we observed a high number of cancelled and restarted EcoOcean runs. Status logs showed that the server often perceived remote clients as having become unresponsive and repeatedly rescheduled their jobs, which we traced back to crucial framework status messages getting intermixed with, and delayed by, the slow transfer of large input and output data files. Swapping over to two separate cloud storage providers (Sync for framework communication, and Dropbox for bulk data transfer), with each cloud storage provider operating on different folders, solved the issue. An important piece of advice was provided by the Dropbox development team, who recommended using the same cloud provider account on the server and all clients to avoid soaring data usage across accounts. Additionally, in rare cases cloud storage providers may stop synchronising which, if this were to occur at the server, stops the framework from working. Some clever coding in the future can detect a hanging cloud provider and restart it.

Although the use of cloud storage providers does demonstrate that a framework can be constructed from the most basic technologies, if faster and more streamlined communication protocols can be used one should not hesitate to embrace those with fervour. For this, the framework is of modular design and already hosts a number of faster and more reliable data communication protocols that require some IT skills and network management authority to configure. To avoid any kind of unnecessary complexity, the use of cloud providers was therefore ideal to showcase the framework.

To demonstrate that the run framework can be OS agnostic, our setup included one Linux computer among five Windows computers. We were able to make this setup work as 1) both the framework and the EcoOcean execution wrapper were written in .NET Standard which natively run on both OS-es; and 2) we were able to fully handle typical OS incompatibilities in our code by enforcing strict data handing conventions. However, as the framework is intended dispatching workloads that rely on any scientific software, framework operators may find that mixed OS family deployments may be very complicated to setup and operate. We recommend that these should be avoided at all costs, and if mixed OS family setups cannot be avoided, we surmise that it may be technically easiest to containerize (Bentaleb et al., 2022) the framework server and clients to the same operating system across available hardware.

We consider the framework that we present here a rough proof of concept that needs to improve in terms of usability, stability and security. In terms of usability, the framework currently offers only bare basic troubleshooting features, collecting execution and error logs in the formats produced natively by the software executed by the user. There are plenty of inspirational methodologies in existence that can be easily adopted (Kandan et al., 2020) to find and understand errors. Additionally, by only collecting software logs but not observing the state of the ecosystem in conjunction, faults caused by the operating system may be presently very hard to identify.

In terms of stability, there is a significant risk of running a framework and launching tasks directly on a target operating system. During the execution of a workload, launched programs may allocate more than their fair share of available resources (such as available processors, runtime memory and disk space). In the worst case, badly behaving programs can crash an operating system. For those with the means and know-how, it would make sense to execute framework clients on virtual machines or containers. These technologies primarily shield the underlying operating system from badly behaving software. Containers increasingly replace virtual machines due to their ability to control resource allocation (Herbein et al., 2016) and to load-balance the use of system resources (e.g., Hota et al., 2019). Such features ensure smoothly flowing executions of hosted software. The framework - or the very concept of the framework - can be easily adapted to execute workloads across more stable environments.

In terms of security, the idea of remote execution of software is generally not encouraged in the world of computing. For this, framework activity must be shielded via secure user authentication and industry standard encryption of all data transferred (e.g., Papadogiannaki & Ioannidis, 2021).

5.2 Case study application

Aside from demonstrating the utility of the framework, the case study also aimed to investigate whether uncertainty within Earth System Models, in the form of retrospective predictions, has the potential to significantly affect the output of a Marine Ecosystem Model as a step towards improving the predictability of MEMs. The brief conclusion is: that depends.

For the functional groups and areas that were explored here, the impact of fishing overwhelmed the impact of ESM uncertainty on EcoOcean results; the natural variability represented in retrospective predictions played a lesser role in affecting EcoOcean outcomes than historical fisheries. For this case study, we did not re-validate EcoOcean's ability to replicate reconstructed catches when driven by ESMs EC-Earth3-CC and NorCPM1. A comprehensive re-validation will be the subject of the oncoming ISIMIP3a simulations (Blanchard et al., 2023). Follow-up work could even consider uncertainty in reconstructed fishing effort. However, these coarse results underscore that effectively managed oceans should prioritise sustainable fisheries practices (e.g., Maury et al., 2017).

The use of time series to reduce data volumes analysed was computationally and storage-wise efficient, but this simplification risks losing important variability in heterogeneous and large areas. As our results showed, aggregating across the entire southern Atlantic obscured trends that became clear when assessing the western and eastern parts of the basin in separation. Future work should explore how to meaningfully measure the sensitivity and performance of a MEM with regards to selecting meaningful regions that are small enough to capture relevant dynamics, and large enough to facilitate speedy analysis. Regional analysis can focus on areas with ecological, geophysical or environmental similarity (e.g., marine ecoregions; Spalding et al., 2007) or other classes of ecoregions (see Rubbens et al. 2023 and references therein). Time periods for comparison should be carefully selected around known events and regime shifts, and possibly even known effects of seasonality (e.g., Lloret-Lloret et al., 2022) and time-delayed teleconnections (e.g., Gómara et al., 2021; Lehodey et al., 2020). We performed a limited time series analysis using only five simple statistics, but there are plenty of other MEM-specific skill metrics suggested in the literature (e.g., Bennett et al., 2013; Hipsey et al., 2020; Kempf et al., 2023; Stow et al., 2009) that could be put to the test. Additionally, advanced vectorization (Quislant et al., 2022) seems to offer significant potential for analysing spatiotemporal MEM output to overcome the limitations of using predefined - and possibly poorly chosen - regions.

As Coll et al. (2020) already identified, driving MEM dynamics with alternative EMSs can come with huge uncertainty, too. The two ESMs included here differed significantly in their approach to representing past environmental conditions. EC-Earth3-CC historical environmental conditions were available from 1980 onwards, starting 30 years later than the NorCPM1 historical data. EcoOcean is calibrated for 1950; and to amend the gap in driver data for EC-Earth3-CC, we applied 1980 driver data for the 1950-1980 period, thus ensuring that EcoOcean had a much more stable spin-up period than when driven by NorCPM1 data. On the other hand, NorCPM1 offered only one phytoplankton group whereas EC-Earth3-CC offered two; different resolutions in the phytoplankton data also meant that both ESMs differently affected food availability to the global food web. For the three sub-ocean basins and six functional groups explored here, EcoOcean showed similar trends under fished and hypothetically non-fished oceans when driven by either ESM, but the trends greatly differed in magnitude depending which ESM was used to drive the environmental conditions for the MEM.

In order to better understand why EcoOcean behaves the way it does, and to quantify if ESM uncertainty has the potential to improve the predictability of EcoOcean, a systematic exploration of attribution is needed to quantify which MEM components are sensitive to which aspects of ESM uncertainty. This could be explored by running different ESM/MEM experiments where ESM internal variability is systematically applied to isolated drivers whilst measuring the impact on MEM output (e.g., Heneghan et al., 2021), and whilst properly validating MEM output against available observations (such as regional trends of species biomasses, regional catch statistics, and global reconstructed fisheries catches). This would also require quantifying the relative importance of other types of uncertainty related to, for instance, trophic structure of the food web and deployed ecological hypotheses (e.g., Coll et al., 2020).

5.3 Future challenges

Up to now, understanding and improving the behaviour of MEMs has been largely a manual process of tweaking model settings guided by intuition and analysing model output (Pethybridge et al., 2019). The framework that we developed here will be the starting point for exploring the effectiveness of proposed skill metrics (Olsen et al., 2016; Payne et al., 2016), validation frameworks (Hipsey et al., 2020) and evaluation protocols (Planque et al., 2022); for assessing various types of uncertainty; and on the long-term, MEM calibration capabilities.

In terms of validation, complex spatial-temporal models are mostly validated by correlating model output with observations (Pethybridge et al., 2019; Spence et al., 2021). However, to ensure that a MEM produces results for the correct reasons, validation should also consider the internal state of a MEM while it executes (e.g., Hipsey et al., 2020; Steenbeek et al., 2021). Indicators of ecosystem dynamics (e.g., Network analysis; Ulanowicz, 2004) and measures of ecological expectations (PREBAL; Link, 2010) can be complemented with assessments of internal state variables related to species displacement, predator/prey overlap, changing environmental conditions and the presence of anthropogenic pressures can capture whether a MEM produces output for the correct reasons, and can provide modellers with valuable insight in the behaviour of their MEMs.

In terms of systematic uncertainty assessments, the multi-run framework provides a foundation for systematically combining parametric uncertainty assessments (e.g., Steenbeek et al., 2018; Vilas et al., 2023) with other forms of uncertainty related to model structure (e.g., Coll et al., 2020; Heneghan et al., 2021), Initialization and Internal variability uncertainty (this case study) and scenario uncertainty (e.g., de Mutsert et al., 2021; Lotze et al., 2019; Schewe et al., 2019; Tittensor et al., 2021).

Combining uncertainty assessments with validation strategies that also consider state variables, modellers can systematically disentangle a model's strengths and weaknesses in search of better model calibrations. Here lies the next big challenge for the global modelling community: to work towards (semi-)automated calibration of spatial-temporal MEMs (e.g., Vilas et al., 2023). By now allowing any modelling group to mass-execute their models systematically across available hardware, the framework can serve as a scaffolding for orchestrating the great number of runs required, which will involve some form of looped and MEM-specific sensitivity testing, parameter estimation and validation scheme.

Here may lie an opportunity for Machine Learning approaches that are increasingly applied to marine ecology (Rubbens et al., 2023). While properly designed statistical approaches can distinguish acceptable from unacceptable ecosystem trends for specific marine ecosystem models, ML approaches can perhaps expand this understanding to infer full food web dynamics from changing environmental conditions, species distributions and fisheries. Following promising work by Trifonova et al. (2017) and Uusitalo et al. (2018), we hope that ML approaches can, one day, assist in the search for more representative parameterizations of complex and mechanistic ecosystem models. A framework such as ours will be essential to mass-execute and perturb MEMs to generate the training datasets needed, and may be able to act as a foundation for ML-assisted MEM calibration.

Our multi-run framework is no silver bullet. With the expansion of computational capabilities also comes the responsibility of using these capabilities wisely. The paradigm "garbage in, garbage out" (A. J. Smith, 1994) is more relevant than ever when scaling up complex model simulations. It is equally useful to utilise the computational capabilities of distributed run frameworks wisely. To avoid wasteful brute-force approaches, one could turn to the use of short press perturbations to identify the most sensitive parameters (Pantus, 2007) and hence dramatically reduce the number of simulations that are really necessary to attain better insights in the workings of complex and mechanistic marine ecosystem models.

Wrapping up

Remote execution frameworks are nothing new, and industry standards greatly surpass the framework described here in all aspects. Our framework shares a number of key principles with Slurm (Yoo et al., 2003), a much more robust and mature, but also much more complicated and technically demanding framework to install and operate. Our framework achieves distributed computing capabilities with the simplest of software components, scaling up desktop workflows, across mundane hardware, without the need for IT skills or programming. That, in itself, is a breakthrough achievement that we hope the global modelling community will build on to make ecosystem modelling operational, for anyone.

The most significant benefit of the framework that we have built here is a full separation of technique (e.g., the technical challenges of repeatedly executing a MEM) from application (e.g., the purpose that the MEM is repeatedly executed for). This allows modellers to just focus on formulating large-scale scientific workloads that the framework then distributes across any available hardware. However, the most significant value of the framework prototype presented here are the ideas within. The simple client/server architecture can be deployed across any hardware configuration: across desktops, virtual machines, docker containers, web servers, and High-Performance Clusters. Any new deployment may require adapting or entirely rewriting framework components to fully utilise hardware capabilities, and to cater to related security and technical constraints. Tech-savvy users may opt to rewrite the multi-run framework in an existing workflow environment. We set out to breach the stigma that MEMs cannot be easily executed systematically; it is now up to the global modelling community to take our ideas further to make the process of ecosystem modelling operational.

The framework presented here is a mere first but important step towards making the process of marine ecosystem modelling more operational. By applying the framework to a global available MEM, we illustrate how it can be useful and how it can be applied to improve our understanding of uncertainty components of complex modelling frameworks, thus opening the door for scientific management breakthroughs.

Acknowledgments

JS acknowledges institutional support from the Polytechnic University of Catalunya (UPC), Barcelona, Spain. RB, VC, MC, EE, NK, PO, DR and JS acknowledge the European Union's Horizon 2020 research and innovation programme under grant agreement N◦ 817578 (TRIATLAS). VC and JS acknowledge the European Commission's Horizon 2020 Research and Innovation programme under grant agreement N◦ 101000302 (EcoScope). JS and MC acknowledge the European Commission's Horizon 2020 Research and Innovation programme under grant agreement N◦ 101060072 (ACTNOW) and the Spanish Ministry of Science and Innovation grant agreement N◦ PID2020-118097RB-I00 (ProOceans). MC acknowledges institutional support of the 'Severo Ochoa Centre of Excellence' accreditation (CEX2019-000928-S). JS, MC and NK conceived the study. NK and DR provided and described the NorESM1 data; RB, EE and PO provided and described EC-Earth3-CC data. MC, JS and VC contributed with EcoOcean. JS developed the MEM run framework and made the ESM data operative with the infinite patience of PO. JS wrote the MS and executed the EcoOcean simulations. VC, LJM, RH, MP and AF analysed the data. All authors contributed to crafting the final paper.

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Open Research

The source code to the MEM multi-run framework prototype is available on GitHub via https://github.com/EwEDevTeam/MEM-Multi-Run-Framework. The code is written in .NET 6.0. can be compliled via Microsoft Visual Studio. The ESM driver data for this MS was generated for the TRIATLAS project and is available via the network of four Earth System Grid

Federation (ESGF) CMIP6 Climate Data portals (USA: https://aims2.llnl.gov/; France: https://esgf-node.ipsl.upmc.fr/search/cmip6-ipsl/; Germany: https://esgf-data.dkrz.de/search/cmip6-dkrz/; UK: https://esgf-index1.ceda.ac.uk/search/cmip6-ceda/). The EcoOcean model is a work in progress, but can be used under a project collaboration by contacting the lead author. The EcoOcean time series output used in this MS is available on Figshare (doi: 10.6084/m9.figshare.24615096). For public and freely available EcoOcean output refer to the ISIMIP data portal at https://data.isimip.org/search/query/EcoOcean/.

References

Årthun, M., Bogstad, B., Daewel, U., Keenlyside, N. S., Sandø, A. B., Schrum, C., & Ottersen, G. (2018). Climate based multi-year predictions of the Barents Sea cod stock. *PLOS ONE*, *13*(10), e0206319. https://doi.org/10.1371/journal.pone.0206319

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

Aumont, O., Ethé, C., Tagliabue, A., Bopp, L., & Gehlen, M. (2015). PISCES-v2: An ocean biogeochemical model for carbon and ecosystem studies. *Geoscientific Model Development*, *8*(8), 2465–2513. https://doi.org/10.5194/gmd-8-2465-2015

Bennett, N. D., Croke, B. F., Guariso, G., Guillaume, J. H., Hamilton, S. H., Jakeman, A. J., Marsili-Libelli, S., Newham, L. T., Norton, J. P., & Perrin, C. (2013). Characterising performance of environmental models. *Environmental Modelling & Software*, *40*, 1–20. https://doi.org/10.1016/j.envsoft.2012.09.011

Bentaleb, O., Belloum, A. S. Z., Sebaa, A., & El-Maouhab, A. (2022). Containerization technologies: Taxonomies, applications and challenges. *The Journal of Supercomputing*, *78*(1), 1144–1181. https://doi.org/10.1007/s11227-021-03914-1

Bentsen, M., Bethke, I., Debernard, J. B., Iversen, T., Kirkevåg, A., Seland, Ø., Drange, H., Roelandt, C., Seierstad, I. A., Hoose, C., & Kristjánsson, J. E. (2013). The Norwegian Earth System Model, NorESM1-M – Part 1: Description and basic evaluation of the physical climate. *Geoscientific Model Development*, *6*(3), 687–720. https://doi.org/10.5194/gmd-6-687-2013

Bethke, I., Wang, Y., Counillon, F., Keenlyside, N., Kimmritz, M., Fransner, F., Samuelsen, A., Langehaug, H., Svendsen, L., Chiu, P.-G., Passos, L., Bentsen, M., Guo, C., Gupta, A., Tjiputra, J., Kirkevåg, A., Olivié, D., Seland, Ø., Solsvik Vågane, J., … Eldevik, T. (2021). NorCPM1 and its contribution to CMIP6 DCPP. *Geoscientific Model Development*, *14*(11), 7073–7116. https://doi.org/10.5194/gmd-14-7073-2021

Bilbao, R., Wild, S., Ortega, P., Acosta-Navarro, J., Arsouze, T., Bretonnière, P.-A., Caron, L.-P., Castrillo, M., Cruz-García, R., & Cvijanovic, I. (2021). Assessment of a full-field initialized decadal climate prediction system with the CMIP6 version of EC-Earth. *Earth System Dynamics*, *12*(1), 173–196. https://doi.org/10.5194/esd-12-173-2021

Blanchard, J., Novaglio, C., Tittensor, D. P., & Forestier, R. (2023). *GitHub—Fish-MIP ISIMIP3a protocol* (FInal). https://github.com/Fish-MIP/FishMIP_2022_3a_Protocol

Boer, G. J., Smith, D. M., Cassou, C., Doblas-Reyes, F., Danabasoglu, G., Kirtman, B., Kushnir, Y., Kimoto, M., Meehl, G. A., Msadek, R., Mueller, W. A., Taylor, K. E., Zwiers, F., Rixen, M., Ruprich-Robert, Y., & Eade, R. (2016). The Decadal Climate Prediction Project (DCPP) contribution to CMIP6. *Geoscientific Model Development*, *9*(10), 3751–3777. https://doi.org/10.5194/gmd-9-3751-2016

Christensen, V., Coll, M., Buszowski, J., Cheung, W. W. L., Frölicher, T., Steenbeek, J., Stock, C. A., Watson, R., & Walters, C. J. (2015). The global ocean is an ecosystem: Simulating marine life and fisheries. *Global Ecology and Biogeography*, *24*(5), 507–517. https://doi.org/10.1111/geb.12281

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*(2–4), 109–139. https://doi.org/10.1016/j.ecolmodel.2003.09.003

Coll, M., Steenbeek, J., Pennino, M. G., Buszowski, J., Kaschner, K., Lotze, H. K., Rousseau, Y., Tittensor, D. P., Walters, C., Watson, R. A., & Christensen, V. (2020). Advancing global ecological modelling capabilities to simulate future trajectories of change in marine ecosystems. *Frontiers in Marine Science*, *7*. https://doi.org/10.3389/fmars.2020.567877

Collins, M. (2002). Climate predictability on interannual to decadal time scales: The initial value problem. *Climate Dynamics*, *19*(8), 671–692. https://doi.org/10.1007/s00382-002-0254-8

Craig, A. P., Vertenstein, M., & Jacob, R. (2012). A new flexible coupler for earth system modeling developed for CCSM4 and CESM1. *The International Journal of High Performance Computing Applications*, *26*(1), 31–42. https://doi.org/10.1177/1094342011428141

Craig, A., Valcke, S., & Coquart, L. (2017). Development and performance of a new version of the OASIS coupler, OASIS3-MCT_3.0. *Geoscientific Model Development*, *10*(9), 3297–3308. https://doi.org/10.5194/gmd-10-3297-2017

Curcin, V., Ghanem, M., & Guo, Y. (2010). The design and implementation of a workflow analysis tool. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *368*(1926), 4193–4208. https://doi.org/10.1098/rsta.2010.0157

de Mutsert, K., Lewis, K. A., White, E. D., & Buszowski, J. (2021). End-to-End Modeling Reveals Species-Specific Effects of Large-Scale Coastal Restoration on Living Resources Facing Climate Change. *Frontiers in Marine Science*, *8*. https://doi.org/10.3389/fmars.2021.624532

Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arsouze, T., Bergman, T., Bernardello, R., Boussetta, S., Caron, L.-P., Carver, G., Castrillo, M., Catalano, F., Cvijanovic, I., Davini, P., Dekker, E., Doblas-Reyes, F. J., Docquier, D., Echevarria, P., Fladrich, U., … Zhang, Q. (2022). The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6. *Geoscientific Model Development*, *15*(7), 2973–3020. https://doi.org/10.5194/gmd-15-2973-2022

Dropbox.com. (2023, June 25). *How to fix Dropbox not syncing and other issues*. https://help.dropbox.com/sync/files-not-syncing

FAO. (2015). *FAO Statistical Areas for Fishery Purposes*. FAO Map Catalog. https://www.fao.org/3/az126e/az126e.pdf

FAO. (2020). *Hydrological basins of the world*. FAO Catalog. https://data.apps.fao.org/catalog/iso/f2615a41-6383-4aa4-aa21-743330eb03ae

Farber, D. J. (1970). *A distributed computer system*. https://escholarship.org/uc/item/4h55t1ww

Fransner, F., Counillon, F., Bethke, I., Tjiputra, J., Samuelsen, A., Nummelin, A., & Olsen, A. (2020). Ocean Biogeochemical Predictions—Initialization and Limits of Predictability. *Frontiers in Marine Science*, *7*. https://doi.org/10.3389/fmars.2020.00386

Gent, P. R., Yeager, S. G., Neale, R. B., Levis, S., & Bailey, D. A. (2010). Improvements in a half degree atmosphere/land version of the CCSM. *Climate Dynamics*, *34*(6), 819–833. https://doi.org/10.1007/s00382-009-0614-8

Gómara, I., Rodríguez-Fonseca, B., Mohino, E., Losada, T., Polo, I., & Coll, M. (2021). Skillful prediction of tropical Pacific fisheries provided by Atlantic Niños. *Environmental Research Letters*, *16*(5), 054066. https://doi.org/10.1088/1748-9326/abfa4d

Halpern, B. S., Frazier, M., Afflerbach, J., Lowndes, J. S., Micheli, F., O'Hara, C., Scarborough, C., & Selkoe, K. A. (2019). Recent pace of change in human impact on the world's ocean. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-47201-9

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*, 102659. https://doi.org/10.1016/j.pocean.2021.102659

Herbein, S., Dusia, A., Landwehr, A., McDaniel, S., Monsalve, J., Yang, Y., Seelam, S. R., & Taufer, M. (2016). Resource Management for Running HPC Applications in Container Clouds. In J. M. Kunkel, P. Balaji, & J. Dongarra (Eds.), *High Performance Computing* (pp. 261–278). Springer International Publishing. https://doi.org/10.1007/978-3-319-41321-1_14

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

Hota, A., Mohapatra, S., & Mohanty, S. (2019). Survey of different load balancing approach-based algorithms in cloud computing: A comprehensive review. *Computational Intelligence in Data Mining: Proceedings of the International Conference on CIDM 2017*, 99–110.

Ilyina, T., Li, H., Spring, A., Müller, W. A., Bopp, L., Chikamoto, M. O., Danabasoglu, G., Dobrynin, M., Dunne, J., Fransner, F., Friedlingstein, P., Lee, W., Lovenduski, N. S., Merryfield, W. j., Mignot, J., Park, J. y., Séférian, R., Sospedra-Alfonso, R., Watanabe, M., & Yeager, S. (2021). Predictable Variations of the Carbon Sinks and Atmospheric CO2 Growth in a Multi-Model Framework. *Geophysical Research Letters*, *48*(6), e2020GL090695. https://doi.org/10.1029/2020GL090695

Iversen, T., Bentsen, M., Bethke, I., Debernard, J. B., Kirkevåg, A., Seland, Ø., Drange, H., Kristjansson, J. E., Medhaug, I., Sand, M., & Seierstad, I. A. (2013). The Norwegian Earth System Model, NorESM1-M – Part 2: Climate response and scenario projections. *Geoscientific Model Development*, *6*(2), 389–415. https://doi.org/10.5194/gmd-6-389-2013

Jeffery, A., Howard, H., & Mortier, R. (2021). Rearchitecting Kubernetes for the Edge. *Proceedings of the 4th International Workshop on Edge Systems, Analytics and Networking*, 7–12. https://doi.org/10.1145/3434770.3459730

Jones, A. K., & Schwans, K. (1979). Task forces: Distributed software for solving problems of substantial size. *Proceedings of the 4th International Conference on Software Engineering*, 315–330.

Kandan, R., Khalid, M. F., Ismail, B. I., Mohd Mydin, M. N., Ahmad, H., & Hoe, O. H. (2020). A Generic Log Analyzer for automated troubleshooting in container orchestration system. *2020 8th International Conference on Information Technology and Multimedia (ICIMU)*, 267–270. https://doi.org/10.1109/ICIMU49871.2020.9243558

Keenlyside, N. S., Latif, M., Jungclaus, J., Kornblueh, L., & Roeckner, E. (2008). Advancing decadal-scale climate prediction in the North Atlantic sector. *Nature*, *453*(7191), Article 7191. https://doi.org/10.1038/nature06921

Kempf, A., Spence, M. A., Lehuta, S., Trijoulet, V., Bartolino, V., Villanueva, M. C., & Gaichas, S. K. (2023). Skill assessment of models relevant for the implementation of ecosystem-based fisheries management. *Fisheries Research*, *268*, 106845. https://doi.org/10.1016/j.fishres.2023.106845

Kirkevåg, A., Iversen, T., Seland, Ø., Hoose, C., Kristjánsson, J. E., Struthers, H., Ekman, A. M. L., Ghan, S., Griesfeller, J., Nilsson, E. D., & Schulz, M. (2013). Aerosol–climate interactions in the Norwegian Earth System Model – NorESM1-M. *Geoscientific Model Development*, *6*(1), 207–244. https://doi.org/10.5194/gmd-6-207-2013

Lehodey, P., Bertrand, A., Hobday, A. J., Kiyofuji, H., McClatchie, S., Menkès, C. E., Pilling, G., Polovina, J., & Tommasi, D. (2020). ENSO Impact on Marine Fisheries and Ecosystems. In *El Niño Southern Oscillation in a Changing Climate* (pp. 429–451). American Geophysical Union (AGU). https://doi.org/10.1002/9781119548164.ch19

Li, Y., Zhang, Y., & Dai, X. (2016). Trophic interactions among pelagic sharks and large predatory teleosts in the northeast central Pacific. *Journal of Experimental Marine Biology and Ecology*, *483*, 97–103.

Link, J. S. (2010). *Ecosystem-based fisheries management: Confronting tradeoffs*. Cambridge University Press.

Lloret-Lloret, E., Albo-Puigserver, M., Giménez, J., Navarro, J., Pennino, M. G., Steenbeek, J., Bellido, J. M., & Coll, M. (2022). Small pelagic fish fitness relates to local environmental conditions and trophic variables. *Progress in Oceanography*, *202*, 102745. https://doi.org/10.1016/j.pocean.2022.102745

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*, 201900194. https://doi.org/10.1073/pnas.1900194116

Ludescher, T., Feilhauer, T., & Brezany, P. (2013). Cloud-Based Code Execution Framework for scientific problem solving environments. *Journal of Cloud Computing: Advances, Systems and Applications*, *2*(1), 11. https://doi.org/10.1186/2192-113X-2-11

Madec, G. & the NEMO team. (2023). *NEMO Ocean Engine Reference Manual*. https://doi.org/10.5281/zenodo.8167700

Maier-Reimer, E. (1993). Geochemical cycles in an ocean general circulation model. Preindustrial tracer distributions. *Global Biogeochemical Cycles*, *7*(3), 645–677. https://doi.org/10.1029/93GB01355

Maier-Reimer, E., Kriest, I., Segschneider, J., & Wetzel, P. (2005). The HAMburg Ocean Carbon Cycle Model HAMOCC5.1—Technical Description Release 1.1 [Application/pdf]. *Berichte Zur Erdsystemforschung, 14*. https://doi.org/10.17617/2.994983

Maleki, M. (2022). Exit status in shell scripting. *Developers Ultimate Guide: Linux Bash Scripting*.

Maury, O., Campling, L., Arrizabalaga, H., Aumont, O., Bopp, L., Merino, G., Squires, D., Cheung, W., Goujon, M., Guivarch, C., Lefort, S., Marsac, F., Monteagudo, P., Murtugudde, R., Österblom, H., Pulvenis, J. F., Ye, Y., & van Ruijven, B. J. (2017). From shared socio-economic pathways (SSPs) to oceanic system pathways (OSPs): Building policy-relevant scenarios for global oceanic ecosystems and fisheries. *Global Environmental Change*, *45*, 203–216. https://doi.org/10.1016/j.gloenvcha.2017.06.007

Meehl, G. A., Goddard, L., Murphy, J., Stouffer, R. J., Boer, G., Danabasoglu, G., Dixon, K., Giorgetta, M. A., Greene, A. M., Hawkins, E., Hegerl, G., Karoly, D., Keenlyside, N., Kimoto, M., Kirtman, B., Navarra, A., Pulwarty, R., Smith, D., Stammer, D., & Stockdale, T. (2009). Decadal Prediction: Can It Be Skillful? *Bulletin of the American Meteorological Society*, *90*(10), 1467–1486. https://doi.org/10.1175/2009BAMS2778.1

Meehl, G. A., Richter, J. H., Teng, H., Capotondi, A., Cobb, K., Doblas-Reyes, F., Donat, M. G., England, M. H., Fyfe, J. C., & Han, W. (2021). Initialized Earth System prediction from subseasonal to decadal timescales. *Nature Reviews Earth & Environment*, *2*(5), 340–357.

Nadeau, D., Doutriaux, C., & Taylor, K. (2018). *Climate Model Output Rewriter (CMOR version 3.3)*. Tech. rep., Lawrence Livermore National Laboratory. https://esgf.llnl.gov/media/2015-F2F/Presentations/Climate-Model-Output-Rewrite-CMOR.pdf

Olsen, E., Fay, G., Gaichas, S., Gamble, R., Lucey, S., & Link, J. S. (2016). Ecosystem Model Skill Assessment. Yes We Can! *PLOS ONE*, *11*(1), e0146467. https://doi.org/10.1371/journal.pone.0146467

Pal, C.-V., Leon, F., Paprzycki, M., & Ganzha, M. (2020). A review of platforms for the development of agent systems. *arXiv Preprint arXiv:2007.08961*.

Pantus, F. (2007). *Sensitivity analysis for complex ecosystem models* [University of Queensland, School of Physical Sciences]. https://espace.library.uq.edu.au/view/UQ:158483

Papadogiannaki, E., & Ioannidis, S. (2021). A Survey on Encrypted Network Traffic Analysis Applications, Techniques, and Countermeasures. *ACM Computing Surveys*, *54*(6), 123:1-123:35. https://doi.org/10.1145/3457904

Park, J.-Y., Stock, C. A., Dunne, J. P., Yang, X., & Rosati, A. (2019). Seasonal to multiannual marine ecosystem prediction with a global Earth system model. *Science*, *365*(6450), 284–288. https://doi.org/10.1126/science.aav6634

Payne, M. R., Barange, M., Cheung, W. W., MacKenzie, B. R., Batchelder, H. P., Cormon, X., Eddy, T. D., Fernandes, J. A., Hollowed, A. B., & Jones, M. C. (2016). Uncertainties in projecting climate-change impacts in marine ecosystems. *ICES Journal of Marine Science*, *73*(5), 1272–1282. https://doi.org/10.1093/icesjms/fsv231

Payne, M. R., Danabasoglu, G., Keenlyside, N., Matei, D., Miesner, A. K., Yang, S., & Yeager, S. G. (2022). Skilful decadal-scale prediction of fish habitat and distribution shifts. *Nature Communications*, *13*(1), Article 1. https://doi.org/10.1038/s41467-022-30280-0

Pethybridge, H. R., Weijerman, M., Perrymann, H., Audzijonyte, A., Porobic, J., McGregor, V., Girardin, R., Bulman, C., Ortega-Cisneros, K., & Sinerchia, M. (2019). Calibrating process-based marine ecosystem models: An example case using Atlantis. *Ecological Modelling*, *412*, 108822. https://doi.org/10.1016/j.ecolmodel.2019.108822

Planque, B., Aarflot, J. M., Buttay, L., Carroll, J., Fransner, F., Hansen, C., Husson, B., Langangen, Ø., Lindstrøm, U., Pedersen, T., Primicerio, R., Sivel, E., Skogen, M. D., Strombom, E., Stige, L. C., Varpe, Ø., & Yoccoz, N. G. (2022). A standard protocol for describing the evaluation of ecological models. *Ecological Modelling*, *471*, 110059. https://doi.org/10.1016/j.ecolmodel.2022.110059

Pohlmann, H., Jungclaus, J. H., Köhl, A., Stammer, D., & Marotzke, J. (2009). Initializing Decadal Climate Predictions with the GECCO Oceanic Synthesis: Effects on the North Atlantic. *Journal of Climate*, *22*(14), 3926–3938. https://doi.org/10.1175/2009JCLI2535.1

Pörtner, H.-O., Karl, D. M., Boyd, P. W., Cheung, W., Lluch-Cota, S. E., Nojiri, Y., Schmidt, D. N., Zavialov, P. O., Alheit, J., Aristegui, J., Armstrong, C., Beaugrand, G., Belkovich, V., Bowler, C., Brewer, P., Church, M., Cooley, S. R., del Monte-Luna, P., Edwards, M., … Wittmann, A. C. (2014). Ocean Systems. In *Climate Change 2014: Impacts, Adaptation,* *and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 411–484). Cambridge University Press. https://epic.awi.de/id/eprint/37501/

Pörtner, H.-O., Roberts, D. C., Poloczanska, E. S., Mintenbeck, K., Tignor, M., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., & Möller, V. (2022). *IPCC, 2022: Summary for policymakers*.

Quislant, R., Fernandez, I., Serralvo, E., Gutierrez, E., & Plata, O. (2022). Exploiting vector extensions to accelerate time series analysis. *2022 30th Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP)*, 55–62. https://doi.org/10.1109/PDP55904.2022.00017

Redondo, J. M., Ortin, F., & Cueva, J. M. (2008). Optimizing reflective primitives of dynamic languages. *International Journal of Software Engineering and Knowledge Engineering*, *18*(06), 759–783. https://doi.org/10.1142/S021819400800388X

Rimal, B. P., Jukan, A., Katsaros, D., & Goeleven, Y. (2011). Architectural requirements for cloud computing systems: An enterprise cloud approach. *Journal of Grid Computing*, *9*, 3–26.

Ritchie, D. M. (1984). The UNIX System: A Stream Input-Output System. *AT&T Bell Laboratories Technical Journal*, *63*(8), 1897–1910.

Rousseau, Y., Watson, R. A., Blanchard, J. L., & Fulton, E. A. (2019). Evolution of global marine fishing fleets and the response of fished resources. *Proceedings of the National Academy of Sciences*, 201820344.

Rousset, C., Vancoppenolle, M., Madec, G., Fichefet, T., Flavoni, S., Barthélemy, A., Benshila, R., Chanut, J., Levy, C., Masson, S., & Vivier, F. (2015). The Louvain-La-Neuve sea ice model LIM3.6: Global and regional capabilities. *Geoscientific Model Development*, *8*(10), 2991–3005. https://doi.org/10.5194/gmd-8-2991-2015

Rubbens, P., Brodie, S., Cordier, T., Destro Barcellos, D., Devos, P., Fernandes-Salvador, J. A., Fincham, J. I., Gomes, A., Handegard, N. O., Howell, K., Jamet, C., Kartveit, K. H., Moustahfid, H., Parcerisas, C., Politikos, D., Sauzède, R., Sokolova, M., Uusitalo, L., Van den Bulcke, L., … Irisson, J.-O. (2023). Machine learning in marine ecology: An overview of techniques and applications. *ICES Journal of Marine Science*, *80*(7), 1829–1853. https://doi.org/10.1093/icesjms/fsad100

Schewe, J., Gosling, S., Reyer, C., Zhao, F., Ciais, P., Elliott, J., Francois, L., Huber, V., Lotze, H., & Seneviratne, S. (2019). State-of-the-art global models underestimate impacts from climate extremes. *Nature Communications*.

Schmidt, D., Stal, M., Rohnert, H., & Buschmann, F. (2000). *Pattern-Oriented Software Architecture, Patterns for Concurrent and Networked Objects, Volume 2*. John Wiley & Sons Inc.

Smith, A. J. (1994). The need for measured data in computer system performance analysis or garbage in, garbage out. *Proceedings Eighteenth Annual International Computer Software and Applications Conference (COMPSAC 94)*, 426–431. https://doi.org/10.1109/CMPSAC.1994.342768

Smith, B., Wårlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J., & Zaehle, S. (2014). Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model. *Biogeosciences*, *11*(7), 2027–2054. https://doi.org/10.5194/bg-11-2027-2014

Smith, D. M., Cusack, S., Colman, A. W., Folland, C. K., Harris, G. R., & Murphy, J. M. (2007). Improved Surface Temperature Prediction for the Coming Decade from a Global Climate Model. *Science*, *317*(5839), 796–799. https://doi.org/10.1126/science.1139540

Sospedra-Alfonso, R., Merryfield, W. J., Boer, G. J., Kharin, V. V., Lee, W.-S., Seiler, C., & Christian, J. R. (2021). Decadal climate predictions with the Canadian Earth System Model version 5 (CanESM5). *Geoscientific Model Development*, *14*(11), 6863–6891. https://doi.org/10.5194/gmd-14-6863-2021

Spalding, M. D., Fox, H. E., Allen, G. R., Davidson, N., Ferdaña, Z. A., Finlayson, M., Halpern, B. S., Jorge, M. A., Lombana, A., Lourie, S. A., Martin, K. D., McManus, E., Molnar, J., Recchia, C. A., & Robertson, J. (2007). Marine Ecoregions of the World: A Bioregionalization of Coastal and Shelf Areas. *BioScience*, *57*(7), 573–583. https://doi.org/10.1641/B570707

Spence, M. A., Thorpe, R. B., Blackwell, P. G., Scott, F., Southwell, R., & Blanchard, J. L. (2021). Quantifying uncertainty and dynamical changes in multi-species fishing mortality rates, catches and biomass by combining state-space and size-based multi-species models. *Fish and Fisheries*. https://doi.org/10.1111/faf.12543

Steenbeek, J., Buszowski, J., Chagaris, D., Christensen, V., Coll, M., Fulton, E. A., Katsanevakis, S., Lewis, K. A., Mazaris, A. D., Macias, D., de Mutsert, K., Oldford, G., Pennino, M., Piroddi, C., Romagnoni, G., Serpetti, N., Shin, Y.-J., Spence, M. A., & Stelzenmüller, V. (2021). Making spatial-temporal marine ecosystem modelling better –

A perspective. *Environmental Modelling & Software*, *145*, 105209. https://doi.org/10.1016/j.envsoft.2021.105209

Steenbeek, J., Buszowski, J., Christensen, V., Akoglu, E., Aydin, K., Ellis, N., Felinto, D., Guitton, J., Lucey, S., Kearney, K., Mackinson, S., Pan, M., Platts, M., & Walters, C. J. (2016). Ecopath with Ecosim as a model-building toolbox: Source code capabilities, extensions, and variations. *Ecological Modelling*, *319*, 178–189. https://doi.org/10.1016/j.ecolmodel.2015.06.031

Steenbeek, J., Coll, M., Gurney, L., Mélin, F., Hoepffner, N., Buszowski, J., & Christensen, V. (2013). Bridging the gap between ecosystem modelling tools using geographic information systems: Driving a food-web model with spatial-temporal primary production data. *Ecological Modelling*, *263*(10), 139–151. https://doi.org/10.1016/j.ecolmodel.2013.04.027

Steenbeek, J., Corrales, X., Platts, M., & Coll, M. (2018). Ecosampler: A new approach to assessing parameter uncertainty in Ecopath with Ecosim. *SoftwareX*, *7*(C), 198–204. https://doi.org/10.1016/j.softx.2018.06.004

Stock, A., Murray, C. C., Gregr, E. J., Steenbeek, J., Woodburn, E., Micheli, F., Christensen, V., & Chan, K. M. A. (2023). *Exploring multiple stressor effects with Ecopath, Ecosim, and Ecospace: Research designs, modeling techniques, and future directions*. https://doi.org/10.1016/j.scitotenv.2023.161719

Stow, C. A., Jolliff, J., McGillicuddy Jr, D. J., Doney, S. C., Allen, J. I., Friedrichs, M. A., Rose, K. A., & Wallhead, P. (2009). Skill assessment for coupled biological/physical models of marine systems. *Journal of Marine Systems*, *76*(1–2), 4–15. https://doi.org/10.1016/j.jmarsys.2008.03.011

Sync.com. (2023, June 25). What do I do if Sync is stuck indexing (or not syncing)? *Sync.Com – Help Desk*. https://www.sync.com/help/what-do-i-do-if-sync-is-stuck-indexing-or-not-syncing/

Thorson, J. T. (2019). Forecast skill for predicting distribution shifts: A retrospective experiment for marine fishes in the Eastern Bering Sea. *Fish and Fisheries*, *20*(1), 159–173. https://doi.org/10.1111/faf.12330

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. *Geosci. Model Dev.*, *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*, 1–9. https://doi.org/10.1038/s41558-021-01173-9

Tjiputra, J. F., Assmann, K., Bentsen, M., Bethke, I., Otterå, O. H., Sturm, C., & Heinze, C. (2010). Bergen Earth system model (BCM-C): Model description and regional climate-carbon cycle feedbacks assessment. *Geoscientific Model Development*, *3*(1), 123–141. https://doi.org/10.5194/gmd-3-123-2010

Trifonova, N., Maxwell, D., Pinnegar, J., Kenny, A., & Tucker, A. (2017). Predicting ecosystem responses to changes in fisheries catch, temperature, and primary productivity with a dynamic Bayesian network model. *ICES Journal of Marine Science*, *74*(5), 1334–1343. https://doi.org/10.1093/icesjms/fsw231

Ulanowicz, R. E. (2004). Quantitative methods for ecological network analysis. *Computational Biology and Chemistry*, *28*(5), 321–339. https://doi.org/10.1016/j.compbiolchem.2004.09.001

Uusitalo, L., Tomczak, M. T., Müller-Karulis, B., Putnis, I., Trifonova, N., & Tucker, A. (2018). Hidden variables in a Dynamic Bayesian Network identify ecosystem level change. *Ecological Informatics*, *45*, 9–15. https://doi.org/10.1016/j.ecoinf.2018.03.003

Vertenstein, M., Craig, T., Middleton, A., Feddema, D., & Fischer, C. (2010). *CCSM4.0 User's Guide*. https://www2.cesm.ucar.edu/models/ccsm4.0/ccsm_doc/book1.html

Vilas, D., Buszowski, J., Sagarese, S., Steenbeek, J., Siders, Z., & Chagaris, D. (2023). Evaluating red tide effects on the West Florida Shelf using a spatiotemporal ecosystem modeling framework. *Scientific Reports*, *13*(1), Article 1. https://doi.org/10.1038/s41598-023-29327-z

Vouk, M. A. (2008). Cloud Computing – Issues, Research and Implementations. *Journal of Computing and Information Technology*, *16*(4), 235–246. https://doi.org/10.2498/cit.1001391

Wang, J., Altintas, I., Berkley, C., Gilbert, L., & Jones, M. B. (2008). A High-Level Distributed Execution Framework for Scientific Workflows. *2008 IEEE Fourth International Conference on eScience*, 634–639. https://doi.org/10.1109/eScience.2008.166

Yoo, A. B., Jette, M. A., & Grondona, M. (2003). Slurm: Simple linux utility for resource management. *Job Scheduling Strategies for Parallel Processing: 9th International Workshop, JSSPP 2003, Seattle, WA, USA, June 24, 2003. Revised Paper 9*, 44–60.