Global Biogeochemical Cycles®

RESEARCH ARTICLE
10.1029/2022GB007367

Key Points:

- On the global scale, five of six models that include mesozooplankton perform moderately well with respect to the observations
- We identify an emergent constraint using the mesozooplankton versus chlorophyll relationship that can help constrain zooplankton projections
- More attention needs to be paid to prey preferences, food web structure and temperature sensitivity in addition to existing key parameters

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

C. M. Petrik,
cpetrik@ucsd.edu

Citation:

Petrik, C. M., Luo, J. Y., Heneghan, R. F., Everett, J. D., Harrison, C. S., & Richardson, A. J. (2022). Assessment and constraint of mesozooplankton in CMIP6 Earth system models. *Global Biogeochemical Cycles*, *36*, e2022GB007367. https://doi.org/10.1029/2022GB007367

Received 28 FEB 2022
Accepted 2 NOV 2022

Author Contributions:

Conceptualization: C. M. Petrik, J. Y. Luo, R. F. Heneghan, C. S. Harrison
Data curation: C. M. Petrik, J. Y. Luo, R. F. Heneghan, J. D. Everett, A. J. Richardson
Formal analysis: C. M. Petrik, J. Y. Luo, R. F. Heneghan, J. D. Everett, A. J. Richardson
Investigation: C. M. Petrik, J. Y. Luo, R. F. Heneghan, J. D. Everett, A. J. Richardson
Methodology: C. M. Petrik, J. Y. Luo, R. F. Heneghan, A. J. Richardson

Assessment and Constraint of Mesozooplankton in CMIP6 Earth System Models

C. M. Petrik[1,2], J. Y. Luo[3], R. F. Heneghan[4], J. D. Everett[5,6,7], C. S. Harrison[8], and A. J. Richardson[5,6]

[1]Department of Oceanography, Texas A&M University, College Station, TX, USA, [2]Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA, [3]NOAA Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA, [4]School of Mathematical Sciences, Queensland University of Technology, Brisbane, QLD, Australia, [5]School of Mathematics and Physics, The University of Queensland, St. Lucia, QLD, Australia, [6]Commonwealth Scientific and Industrial Research Organisation (CSIRO) Oceans and Atmosphere, Queensland Biosciences Precinct, St. Lucia, QLD, Australia, [7]Centre for Marine Science and Innovation, The University of New South Wales, Sydney, NSW, Australia, [8]Department of Ocean and Coastal Science, Center for Computation and Technology, Louisiana State University, Baton Rouge, LA, USA

Abstract Although zooplankton play a substantial role in the biological carbon pump and serve as a crucial link between primary producers and higher trophic level consumers, the skillful representation of zooplankton is not often a focus of ocean biogeochemical models. Systematic evaluations of zooplankton in models could improve their representation, but so far, ocean biogeochemical skill assessment of Earth system model (ESM) ensembles have not included zooplankton. Here we use a recently developed global, observationally based map of mesozooplankton biomass to assess the skill of mesozooplankton in six CMIP6 ESMs. We also employ a biome-based assessment of the ability of these models to reproduce the observed relationship between mesozooplankton biomass and surface chlorophyll. The combined analysis found that most models were able to reasonably simulate the large regional variations in mesozooplankton biomass at the global scale. Additionally, three of the ESMs simulated a mesozooplankton-chlorophyll relationship within the observational bounds, which we used as an emergent constraint on future mesozooplankton projections. We highlight where differences in model structure and parameters may give rise to varied mesozooplankton distributions under historic and future conditions, and the resultant wide ensemble spread in projected changes in mesozooplankton biomass. Despite differences, the strength of the mesozooplankton-chlorophyll relationships across all models was related to the projected changes in mesozooplankton biomass globally and in regional biomes. These results suggest that improved observations of mesozooplankton and their relationship to chlorophyll will better constrain projections of climate change impacts on these important animals.

Plain Language Summary Zooplankton are marine animals that have a key role in transferring carbon from the atmosphere deeper into the ocean. They also serve as a crucial link in food chains between microscopic marine plants (phytoplankton) and predators like fish and whales. Researchers have created mathematical representations (models) of the linked processes of the oceans, atmosphere, and land, most of which include zooplankton. Yet how well these models represent zooplankton has not been adequately tested. We compared observations of zooplankton biomass to model estimates. We explored if these models reproduce the observed relationship between zooplankton biomass and chlorophyll concentration, which is useful for assessing how well the models represent predator-prey relationships. Five of six models had similar patterns and comparable average biomasses across the global ocean as the observations. The historic relationship in three models fell within the observed relationship. The strength of the relationships across all models was related to how much zooplankton biomass will decrease with climate change. To improve the representation of zooplankton in models, we need better observations of the relationships between organisms. This would advance estimates of carbon transfer to the deep sea and carbon available to fish, and how they will change with climate change.

1. Introduction

Zooplankton vary over 16 orders of magnitude in size and constitute 40% of total marine biomass (Hatton et al., 2021). They lie in the middle of marine food webs, an essential position linking primary production to

Resources: R. F. Heneghan
Supervision: C. M. Petrik
Visualization: C. M. Petrik, R. F. Heneghan
Writing – original draft: C. M. Petrik, J. Y. Luo, R. F. Heneghan, J. D. Everett, A. J. Richardson
Writing – review & editing: C. M. Petrik, J. Y. Luo, R. F. Heneghan, J. D. Everett, C. S. Harrison, A. J. Richardson

higher trophic levels (Ryther, 1969). Additionally, zooplankton influence the biological carbon pump both indirectly by modulating the proportion of production that is recycled versus exported, and directly by contributing fast-sinking fecal pellets and by vertical migration to depth (Steinberg & Landry, 2017). Ocean biogeochemistry models were historically built to simulate the carbon cycle (Maier-Reimer & Hasselmann, 1987; Sarmiento et al., 1992) and have varying degrees of representation of plankton food web dynamics (Kearney et al., 2021; Séférian et al., 2020). Zooplankton, if present, function more often as a closure term on the phytoplankton, ensuring that phytoplankton biomass does not grow uncontrollably. Zooplankton are less often considered as an ecologically or biogeochemically meaningful variable, yet, the response of phytoplankton is extremely sensitive to the formulation and parameter values of zooplankton (Gentleman & Neuheimer, 2008; Rohr et al., 2022).

Coupled Earth system models (ESMs) used to study global climate change include biogeochemistry within their ocean models because of the outsized role of the ocean in the global carbon cycle (e.g., Friedlingstein et al., 2020) and thus the climate system. Also, plankton food web dynamics from these ESMs are increasingly used in marine ecosystem models as forcings that represent the base of marine food webs (e.g., Harrison et al., 2021; Heneghan et al., 2021; Lotze et al., 2019; Tittensor et al., 2018, 2021). Many of these marine ecosystem models use phytoplankton outputs from ESMs, which have undergone formal skill assessment in terms of chlorophyll *a* or net primary production in comparison to satellite-based products (Bopp et al., 2013; Fu et al., 2015, 2022; Kwiatkowski et al., 2020; Laufkötter et al., 2015; Séférian et al., 2020). However, formal skill assessment of biogeochemistry in ESMs has not yet extended to the zooplankton, although individual models have performed some comparison of simulated zooplankton with observations (Aumont et al., 2018; Stock et al., 2014b; Yool et al., 2021).

Skill assessment of zooplankton is needed for several reasons. First, improving zooplankton representation based on ecological mechanisms and not the tuning of parameters would provide more realistic estimates of phytoplankton biomass and production, export, and the whole carbon cycle (Laufkötter et al., 2015, 2016). As the role of zooplankton in the export of carbon to the deep ocean is one of the largest uncertainties in ESM simulations of the marine carbon cycle, constraining their biomass is important for understanding future climate (Henson et al., 2021). Second, ESM projections demonstrate an amplified response of mesozooplankton biomass and production relative to phytoplankton (Kwiatkowski et al., 2019; Stock et al., 2014a). Third, some marine ecosystem models directly use zooplankton output from ESMs to force higher trophic levels (Maury, 2010; Petrik et al., 2019). Last, if model skill in simulating zooplankton exceeds or has less uncertainty than that of phytoplankton, then it may be a more robust forcing term for models of upper trophic levels.

Zooplankton can be divided into multiple classes that have varying ecosystem functions and observational methods. We focus here on the mesozooplankton (0.2–20 mm; Sieburth et al., 1978) that are prey of higher trophic level animals such as forage fish, seabirds, and whales. Mesozooplankton also are more representative of export production, whereas microzooplankton (<0.2 mm) are an important component of the microbial loop, generating recycled production rather than transferring carbon up the food chain (Buitenhuis et al., 2006). Furthermore, most observations target the mesozooplankton size range, massively underestimating both microzooplankton and macrozooplankton (>20 mm; Wiebe & Benfield, 2003), making it a better analog for comparing against observations than total zooplankton biomass.

An examination of mesozooplankton across ESMs will aid understanding and reduce uncertainty in climate projections, especially as biological variables have larger structural uncertainty than physical variables and are more influenced by future climate scenario uncertainty (Frölicher et al., 2016; Kwiatkowski et al., 2020; Tittensor et al., 2021; Fu et al., 2022). Multimodel ensembles are often used to capture a wide range of structural uncertainty, and standard skill metrics can help to identify skillful models for weighting or constraining ensembles (Brunner et al., 2020; Séférian et al., 2020). Another method for reducing projection uncertainty is the *emergent constraint* method, which uses a combination of a model's match to an observable relationship and the model ensemble itself to constrain future projections (Eyring et al., 2019; Hall & Qu, 2006; Hall et al., 2019; Kwiatkowski et al., 2017). This approach requires two components: (a) a physically interpretable and observable relationship between two quantities that have a degree of variation in the contemporary climate; and (b) an emergent statistical relationship between the first relationship and an Earth System sensitivity that can be quantified across an ensemble of ESMs (Eyring et al., 2019). One of the earliest examples within ocean biogeochemistry is from Kwiatkowski et al. (2017), where the authors identified an observable relationship between net primary production (NPP) and sea surface temperature (SST) anomaly in the Niño 3.4 region, which was also present in ESM pre-industrial

simulations. Furthermore, models with stronger sensitivity of tropical NPP to SST in pre-industrial simulations projected greater declines in NPP in climate change simulations. Because of this tight correlation between the observable relationship and future change, the authors were able to use models' agreement with the observed relationships to constrain future projections of NPP change. We use both the standard skill assessment and emergent constraints approaches here, estimating performance by evaluating the mesozooplankton output from individual ESMs against observations through skill assessment, and constraining future projections of mesozooplankton change in ESMs through an emergent constraints analysis.

In this study, we assess mesozooplankton in ESMs to answer four key questions. (a) How well do ESMs represent mesozooplankton under historic conditions? (b) Are there regions or seasons where ESM representation of mesozooplankton is better or worse? (c) Are ESMs able to reproduce the relationships between mesozooplankton biomass and surface chlorophyll *a* seen in observations? (d) Do the relationships between mesozooplankton biomass and surface chlorophyll *a* under historic conditions characterize the sensitivity of mesozooplankton to climate change across ESMs? Our hope is that through greater assessment of mesozooplankton in ESMs, and biogeochemistry models more generally, we can improve their representation, thus enabling more robust model projections of fisheries and carbon cycling.

2. Methods

2.1. Earth System Models

We examined mesozooplankton biomass simulated by all ESMs in the sixth phase of the Coupled Model Intercomparison Project (CMIP6; Eyring et al., 2016) with a dedicated mesozooplankton group. The CMIP6 zooplankton fields that modeling groups could submit were biomass of microzooplankton, mesozooplankton, and total zooplankton. When model description papers used the terms small zooplankton and large zooplankton, we assumed these were the microzooplankton and mesozooplankton outputs, respectively, unless model descriptions indicated otherwise (e.g., Stock et al., 2014b, 2020). The models with mesozooplankton are CanESM5-CanOE, CMCC-ESM2, CNRM-ESM2-1, GFDL-ESM4, IPSL-CM6A-LR, and UKESM1-0-LL (Table S1 in Supporting Information S1). Other models that participated in CMIP6 either have no explicit zooplankton group or only one zooplankton group. We used the Historic and "very high emissions" (SSP5-8.5) scenario simulations of these models (Allan et al., 2021). Model outputs were downloaded from the ESGF-CoG (CanESM5-CanOE, CMCC-ESM2, CNRM-ESM2-1) and ISIMIP (GFDL-ESM4, IPSL-CM6A-LR, and UKESM1-0-LL) servers. This analysis focused on model outputs of mesozooplankton biomass (zmeso, molC m^{-3}), surface chlorophyll concentration (chlos, kg m^{-3}), sea surface temperature (tos, °C), depth level thickness (thkcello, m), and mixed layer depth (mlotst, m). Note that parentheses give the CMIP6 standard name and units of each variable.

The ocean biogeochemistry (BGC) sub-models of each ESM are described to varying degrees in model description publications (Table S1 in Supporting Information S1), with a comprehensive summary and comparison of their structure performed by Kearney et al. (2021) and a comparison of model skill (excluding zooplankton) by Séférian et al. (2020). There are considerable differences in the formulation of the BGC sub-models of the ESMs, particularly with respect to zooplankton (Table 1). Here we briefly give details of each sub-model relevant to the present study. All BGC sub-models included here simulate interactions among nutrients, phytoplankton, zooplankton, and detritus. Phytoplankton require the uptake of nutrients to fix CO_2 into organic carbon and are then grazed by zooplankton. Metabolic, feeding, and mortality processes of the phytoplankton and zooplankton lead to the production of detritus and remineralization of dissolved nutrients. As these processes are linked, structural and parameter differences in their formulations could directly or indirectly affect zooplankton. For instance, functions for phytoplankton uptake of N, P, Fe, Si, and their light limitation vary widely across BGC sub-models (Kearney et al., 2021), which could result in different phytoplankton dynamics that then impact zooplankton grazers.

There are many other differences across BGC sub-models, including the number and type of plankton groups, predator-prey relationships, feeding functional responses, the temperature dependence of biological rates, and loss terms for zooplankton (Table 1). Each BGC sub-model is different in its formulation and parameterization. However, two ESMs use the same BGC sub-model: CNRM-ESM2-1 and IPSL-CM6A-LR use PISCES2.0 (Aumont et al., 2015), except that CNRM uses PISCES2.0-gas, which resolves dimethylsulfide and nitrous oxide. The CanESM5-CanOE (Christian et al., 2021), PISCES2.0, and UKESM1-0-LL-MEDUSA2.1 (Yool

Global Biogeochemical Cycles

10.1029/2022GB007367

Table 1
Differences in the ESMs Relevant for Mesozooplankton Including the Phytoplankton (P), Zooplankton (Z), and Detritus (D) Groups, Mesozooplankton Prey Types, Grazing Functional Response, Key Grazing Parameters, Temperature Dependence of Rate Processes, and Key Reference

Model	P groups	Z groups	D groups	MesoZ prey	Grazing functional response	Grazing parameters	Temperature dependent processes	Reference
BFMv5.2 (CMCC)	2 NanoP, Diatoms	2 MicroZ, MesoZ	1	MesoZ: Diatoms, MicroZ, MesoZ	Type II (Disk)	MesoZ: Potential specific growth rate = 2.0 d^{-1}, k = 80 mgC m^{-3}[a]	Q_{10} = 2 on all rate processes	Lovato et al. (2022)
CanOE (CAN)	2 SmallP, LargeP	2 SmallZ, LargeZ	2 Small/slow-sinkingD, Large/fast-sinkingD	LargeZ: LargeP, SmallZ	Type II (Ivlev)	LargeZ: $gmax$ = 0.85 d^{-1}, k = 33 mgC m^{-3}[b]	E_a = 37.4 kJ mol^{-1} for P growth and Z respiration. No T dependence on Z grazing nor P and Z mortality	Christian et al. (2021)
COBALTv2 (GFDL)	3 Diazotrophs, SmallP, LargeP	3 SmallZ (<200 µm), MediumZ[d] (200–2,000 µm),	1	MediumZ: LargeP, Diazotrophs, SmallZ; LargeZ: LargeP, Diazotrophs, MediumZ	Type II (Michaelis-Menton), with mild density-dependent prey switching	MediumZ: $Imax$ = 0.57 d^{-1}, k = 99 mgC m^{-3}; LargeZ: Imax = 0.23 d^{-1}, k = 99 mgC m^{-3}[c]; Imax decreases with Z size and k kept constant	Q_{10} = 1.878 for P growth, P mortality, Z grazing, and Z mortality	Stock et al. (2014b, 2020)
MEDUSA2.1 (UK)	2 Non-Diatoms, Diatoms	2 MicroZ, MesoZ	2 Small/slow-sinkingD, Large/fast-sinkingD	MesoZ: Non-Diatoms, Diatoms, MicroZ, SmallD	Type III (Michaelis-Menton)	MesoZ: $gmax$ = 0.5 d^{-1}, k = 24 mgC m^{-3}[e]	Q_{10} = 1.895 for P growth and mortality. No T dependence for Z grazing or Z mortality	Yool et al. (2013, 2021)
PISCES2.0 (CNRM, IPSL)	2 NanoP, Diatoms	2 MicroZ, MesoZ	2 Small/slow-sinkingD, Large/fast-sinkingD	MesoZ: NanoP, Diatoms, SmallD, LargeD	Type II (Michaelis-Menton), with no prey switching	MesoZ: $gmax$ = 0.75 d^{-1}, k = 242 mgC m^{-3}[f]	Q_{10} = 1.895 for P growth; 2.14 for Z grazing and Z mortality. No T dependence for P mortality	Aumont et al. (2015)

Note. "No temperature-dependence" is equivalent to a Q_{10} = 1. $gmax$: maximum grazing rate; $Imax$: maximum ingestion rate; k: half saturation constant for grazing or ingestion rate.

[a]k = Potential specific growth rate/Specific search volume, where Specific search volume = 0.025 m^3 mgC^{-1} d^{-1}. [b]k = −ln (0.5)/a, where a = 0.021 m^3 mgC^{-1} converted from 0.25 (mmol C m^{-3})$^{-1}$. [c]Converted from 1.25 µmol N kg^{-1} assuming Redfield Ratio of 106:16:1 (C:N:P). [d]Mesozooplankton in COBALT combines medium and large zooplankton. [e]Converted from 0.3 mmol N m^{-3} assuming Redfield Ratio of 106:16:1 (C:N:P). [f]Converted from 20 µmol C L^{-1}. All conversions to mgC used mg = µmol × 12.0107 × 1e−3 or mg = nmol × 12.0107 × 1e−6.

et al., 2013, 2021) models all include two phytoplankton (small/nano and large/diatoms), two zooplankton (small/micro, large/meso), and two particulate detritus (small/slow-sinking and large/fast-sinking) groups. The CMCC-ESM2-BFMv5.2 (Lovato et al., 2022) model has one heterotrophic bacteria group, two phytoplankton, two zooplankton, and one particulate detritus term. The GFDL-ESM4-COBALTv2 (Stock et al., 2014b, 2020) model represents bacteria, diazotrophs, small phytoplankton, large phytoplankton, small zooplankton, medium zooplankton, large zooplankton, and particulate detritus. The mesozooplankton biomass provided by COBALTv2 in CMIP6 is the combination of the medium and large zooplankton groups. All models represent losses to mesozooplankton biomass with both linear (constant) and non-linear (density-dependent) mortality terms. Linear terms often represent metabolic losses while density-dependent terms represent predation by higher trophic levels. The density-dependent mortality function is quadratic in CanOE, PISCES2.0, and COBALTv2, while it is a different power function in BFMv5.2 and MEDUSA2.1 (Yool et al., 2011). In PISCES2.0, the linear mortality is increased at low dissolved oxygen concentrations to mimic hypoxic stress. Since the linear mortality term in BFMv5.2 represents senescence, there are additional losses to excretion/egestion, respiration, cannibalism, and an oxygen-dependent linear mortality term. The medium zooplankton of COBALTv2 also experience mortality when eaten by the large zooplankton.

Each BGC sub-model has a different set of prey available to mesozooplankton (Table 1). CanOE has the most restricted diet, with mesozooplankton feeding only on the large phytoplankton and small zooplankton groups. Mesozooplankton in BFMv5.2 prey similarly on the large phytoplankton and small zooplankton, but also perform cannibalism. MEDUSA2.1 has more prey items that include small phytoplankton, large phytoplankton, small zooplankton, and small detritus. PISCES2.0 mesozooplankton have the broadest diet: small phytoplankton, large phytoplankton, small zooplankton, and both sizes of detritus. In COBALTv2, both medium and large zooplankton prey on diazotrophs and large phytoplankton. Medium zooplankton additionally prey on small zooplankton, while large zooplankton also consume the medium zooplankton.

In addition to varied diets, ingestion by zooplankton is modeled with different functional responses. BFMv5.2, CanOE, and COBALTv2 use a Type II functional response, PISCES2.0 uses a Type II response with a threshold, and MEDUSA2.1 uses a Type III sigmoid response. All consider total prey availability when calculating grazing rates. Grazing rates in CanOE are proportional to abundance across all prey types, while BFMv5.2, PISCES2.0, COBALTv2, and MEDUSA2.1 use some form of prey preference. The prey preference in COBALT is dependent on abundance to allow for prey switching, which imparts sigmoidal characteristics on the functional response. In addition to the functional response, grazing rate parameters also differ (Table 1). Ingestion is influenced by zooplankton and phytoplankton nutrient ratios because the zooplankton have fixed stoichiometry in these models, but the phytoplankton either have variable stoichiometry (BFMv5.2, CanOE, PISCES2.0) or fixed stoichiometry that differs from the zooplankton (COBALTv2, MEDUSA2.1). Though models differ in which ratios they consider (C:N:Fe – CanOE; C:N:P – BFMv5.2, COBALTv2; C:N – MEDUSA2.1), nutrients in excess of the zooplankton stoichiometry return to dissolved or particulate pools through egestion, excretion, and respiration. PISCES2.0 additionally uses N:C and Fe:C ratios as an indicator of prey quality, where decreasing ratios reduce growth efficiency.

Probably the greatest diversity in the BGC sub-models is in the temperature dependence of biological rates (Table 1). To impose these relationships, CanOE uses the Arrhenius-Van't Hoff equation, BFMv5.2 employs a non-dimensional Q_{10} function, and COBALTv2 and PISCES2.0 use an Eppley curve (Eppley, 1972). Sub-models differ both with respect to the biological rates with temperature dependence and the strength of this temperature dependence. Temperature dependence is applied to phytoplankton growth rates in all BGC sub-models. PISCES2.0 and COBALTv2 apply temperature dependence to zooplankton grazing rates, with a sensitivity greater than phytoplankton growth in PISCES2.0 and equal sensitivity in COBALTv2. COBALTv2 also applies temperature dependence to phytoplankton and zooplankton loss terms (but not phytoplankton aggregation), while PISCES2.0 only applies it to zooplankton losses. Conversely, CanOE applies temperature dependence to zooplankton respiration rates only. Finally, detritus remineralization is a function of temperature in CanOE, PISCES2.0, and COBALTv2. The strength of this relationship is equal to the phytoplankton growth relationship in COBALTv2 and PISCES2.0, and temperature has a stronger influence on remineralization compared to phytoplankton growth in CanOE. BFMv5.2 uses a constant Q_{10} value for all phytoplankton and zooplankton physiological processes. MEDUSA2.1 does not apply temperature dependence on the zooplankton rates.

2.2. Mesozooplankton Observation-Based Statistical Model

We compared the historical mesozooplankton biomass from ESMs with an observation-based statistical model. However, we first considered comparing zooplankton biomass from ESMs with observations from Moriarty and O'Brien (2013) derived from the valuable COPEPOD database. We decided against using these zooplankton measurements directly because of their unknown biases due to the large variation in how and under what conditions the data were collected, and the standardization used to adjust for these biases in Moriarty and O'Brien (2013). For example, zooplankton biomass was measured on samples using six different methods (carbon, ash free dry weight, dry weight, wet weight, displacement volume, settled volume), and observations were made with dozens of different nets and 43 different mesh sizes, all of which hugely influence the biomass measured (Everett et al., 2017; Postel et al., 2000). However, the conversions used for standardization of the measurement method to carbon were based on a relatively small number of observations restricted in space and time (see O'Brien [2005] for details and references therein). The particular zooplankton community present in a sample, especially the ratio of crustacean to gelatinous zooplankton present, will highly influence these standardizations (Everett et al., 2017). Similarly, the conversion to a mesh size of 333 μm was also based on a small number of measurements restricted in space and time (O'Brien, 2005), despite the particular zooplankton community present in a sample influencing the size (affects extrusion through the net) and mobility (affects net avoidance by zooplankton) and thus the biomass measured. Further, there was no standardization for many other potential biases, including the 65 different net types used, different times of day to account for the effect of diel vertical migration on measured biomass, different seasonal cycles in different areas, or the decline of zooplankton biomass as you sample deeper in the water column. Using simple standardizations based on data that have a plethora of complex biases is challenging.

Instead, to assess the historical mesozooplankton biomass from ESMs, we used a Generalized Linear Mixed Model (GLMM) approach, which is commonly used in ecology to adjust for observation biases using a suite of fixed and random predictors (Bolker et al., 2009). We used the GLMM of Heneghan et al. (2020; hereafter called obsGLMM), which combines data from Moriarty and O'Brien (2013) with additional observations (total $n = 196{,}907$). The obsGLMM adjusts for temporal, spatial, and sampling biases within the data set by including a suite of fixed and random predictors. The obsGLMM included two random effects, one that accounted for gear type that represented the effect on zooplankton biomass of the different capture efficiencies and avoidance of each net, and another that accounted for institution that represented the effect of different ships, how nets might be towed, and any differences in how the samples are processed. Fixed effects account for biases associated with measurement method (with levels for settled volume, displacement volume, wet weight, dry weight, ash free dry weight, carbon and biovolume), mesh size of the different nets used (continuous, 50–1,000 μm), the mean depth of sampling (continuous, 0–1,500 m), bathymetry (0 to >7,000 m), sampling day of year (continuous, all days sampled) to account for seasonality, time of day (continuous, all hours sampled) to adjust for diel vertical migration, and an estimate of the current environmental conditions based on sea surface temperature (continuous, −2 to 32°C) and surface chlorophyll *a* (0.001–10 mg m^{-3}).

For the environmental predictors, monthly climatologies of sea surface temperature and satellite chlorophyll *a* measurements were used because data included times prior to satellite observations. Data were obtained from MODIS-Aqua (Moderate Resolution Imaging Spectroradiometer aboard the Aqua spacecraft, 4 km resolution) averaged over 2002–2016, aggregated to a 1° spatial resolution and accessed via the NASA Giovanni portal (https://giovanni.gsfc.nasa.gov/giovanni/). Bathymetry data for the model were sourced from GEBCO (General Bathymetric Chart of the Oceans; https://www.gebco.net).

The obsGLMM was used to produce global estimates of observed mesozooplankton biomass that have the same spatial resolution as the mesozooplankton output from the ESMs. We assume it estimates mesozooplankton biomass because most net samples capture mesozooplankton predominantly, with few microzooplankton captured because the meshes are too coarse, although some macrozooplankton will be sampled (Wiebe & Benfield, 2003). The obsGLMM estimates the mean mesozooplankton across years as there were insufficient data over time to produce robust multi-annual time series, but the day of year term allows us to estimate a seasonal climatology for zooplankton biomass.

Although, the obsGLMM captures most of the variability in the COPEPOD data, with an R^2 of 91% (82% from fixed effects), there are caveats associated with the obsGLMM. The main one is that the obsGLMM from Heneghan et al. (2020) is an estimate of mean biomass, not an interpolation, and thus smooths over considerable

spatial and temporal variation. The reason that this biomass field does not simply interpolate between observations is because of the many different biases in sampling we have highlighted and the patchy distribution of zooplankton observations in time and space. However, we believe this caveat is outweighed by the strengths of the obsGLMM, which adjusts for sampling biases and different biomass measurement methods and allows global gridded fields based on all available data.

2.3. Other Mesozooplankton Observational Products

Supplementary skill assessments were performed with two other zooplankton datasets: (a) the above-mentioned Moriarty and O'Brien (2013) point observations of zooplankton carbon biomass (mgC m^{-3}, n = 10,117) integrated over the top 200 m (obsMO hereafter), and (b) the Strömberg et al. (2009) empirical model of mesozooplankton biomass (obsSM). The obsMO are the raw observations from the broader COPEPOD database that have undergone some standardization of units (Moriarty & O'Brien, 2013), though it is limited in terms of accounting for conversion and sampling biases compared to the obsGLMM. The obsSM is an observationally based (n = 4,843) trophic transfer model forced with SeaWiFS satellite ocean color data to produce a global product of biomass estimates (Strömberg et al., 2009). The obsSM does not adjust for the many biases in the observational zooplankton data as the obsGLMM does, but both provide a global gridded product with no gaps in space or time. Both products are climatologies, with monthly, seasonal, and annual means available.

2.4. Processing for Model-Observation Comparisons

Observation-based estimates and ESM output of mesozooplankton biomass were made as compatible as possible before quantitative comparison. We constrained estimates of mesozooplankton biomass to the top 200 m of the water column (the epipelagic zone), where densities of phytoplankton and zooplankton are highest. ESM mesozooplankton output in molC m^{-3} was depth-integrated over the top 200 m using the native model grid, converted from molC m^{-2} to mgC m^{-2} using 12.01 gC per mol, and regridded with linear interpolation to a common $1° \times 1°$ grid. Similarly, total mesozooplankton biomass in the top 200 m from the obsGLMM, obsMO, and obsSM were calculated on the same $1° \times 1°$ grid. Few zooplankton data were collected before 1950 and the last addition of new observations was in 2015. For comparison with the observation-based estimates, the 50-year period from 1965 to 2014 was analyzed from the Historical ESM simulations. We computed seasonal and annual climatologies from all 600 months of ESM output by first creating a 12-month climatology, then taking the 3-month mean for each season and 12-month mean for the annual. Prior to averaging, southern hemisphere seasons were standardized to align with the northern hemisphere (e.g., winter corresponds to northern hemisphere December–February and southern hemisphere June–August). Absolute biomasses were $\log_{10}$ transformed before comparison to remove skewness and the effect of outliers.

2.5. Spatial Scale and Biome Definition

The coarse spatial resolution of global ESMs (~25–100 km) in CMIP6 is best suited to simulating large-scale regional differences in the global ocean and is less skillful at accurately simulating quantities at specified times and locations, thereby making point comparisons difficult. Therefore, in addition to making geographic comparisons by grid cell, we analyzed oceanic biomes independently with the expectation that ecosystem processes function similarly in each of these regions. Ocean biomes are usually defined by the biophysical characteristics that drive different ecosystem types: temperature stratification, which affects nutrient delivery, light limitation, and plankton community structure (Behrenfeld & Boss, 2018), and chlorophyll a, an indication of phytoplankton productivity. We used the biomes defined by Stock et al. (2014b), following Banse (1992): (a) LC—low chlorophyll; (b) HCPS—high chlorophyll, permanently stratified; and (c) HCSS – high chlorophyll, seasonally stratified. In Stock et al. (2014b), the distinction between the low and high chlorophyll biomes from observations was based on a chlorophyll a threshold of 0.125 mg chl m^{-3}, and the seasonally and permanently stratified biomes were differentiated by the maximum annual mixed layer depth from the monthly climatology being seasonally above or permanently below 75 m depth. The choice of the three biomes defined by Stock et al. (2014b) represents a conservative choice, as other options range from 7 to 56 biomes or provinces (Longhurst, 1995; Sarmiento et al., 2004).

Historical runs of ESMs often perform relatively poorly against satellite estimates of ocean chlorophyll *a*, with both significant global and regional biases (Fu et al., 2022; Séférian et al., 2020), which presents a challenge to setting the threshold between low and high chlorophyll biomes that accounts for mean biases between the CMIP6 ESMs and chlorophyll *a* observations. Following Stock et al. (2014b), a single chlorophyll *a* threshold was used for each model to define the LC region. We calculated the total ocean surface area delineated by a LC threshold of 0.125 mg chl m^{-3} in satellite-based chlorophyll *a* estimates, roughly 35% of the ocean surface. Then a chlorophyll *a* value from each ESM that would give the same subtropical gyre area as the observational data was determined from the Historical 50-year mean. Thus, the LC threshold is defined as the lowest 35th percentile of chlorophyll *a* weighted by area, limited to the area between 45°N and 45°S. Polar areas were excluded due to poor satellite coverage. Though the chlorophyll *a* threshold value varies by ESM (Figure S2 in Supporting Information S1), each threshold delineates a biome with plankton dynamics consistent with oligotrophic gyres. This historical LC threshold was used to establish the biomes for each ESM under the SSP5-8.5 scenario as well.

For the delineation between permanently versus seasonally stratified high chlorophyll biomes, we followed Stock et al. (2014b) by using the mixed layer depth from an annual climatology of the Historical 50-year time period. This yielded biomes consistent with upwelling zones (permanently stratified) and high latitude seasonal seas (seasonally stratified).

The size and location of biomes varied across models, both historically and in future projections (Figure S2 in Supporting Information S1). There were some discrepancies, particularly in the polar regions, as the use of the mixed layer depth threshold resulted in some portions of the Arctic to be labeled as High Chlorophyll, Permanently Stratified (HCPS) in some models. Still, the choice of using mixed layer depth to define permanently versus seasonally stratified reflects the availability of CMIP6 model outputs; mean irradiance in the mixed layer would be a better variable, but it is not available from all models. Additionally, by definition, the LC biomes represented 35% of the area between 45°N and 45°S, though many models had chlorophyll concentrations below the LC threshold in the Arctic. The HCSS biome appears the most consistent across ESMs, while the HCPS the most variable, likely due to the Arctic influence in some models. Upwelling regions tend to be classified as HCPS, thus the majority of differences in HCPS biomes reflects the ability of the physical ocean model in each ESM to simulate upwelling at the coarse spatial scale. Notably, the HCPS biome in all ESMs shrinks in future simulations, while the LC biome expands.

2.6. Analyses

2.6.1. Comparisons of Observed Versus Modeled Mesozooplankton

Initial comparisons of the ESMs were made against the obsGLMM estimate of mesozooplankton carbon biomass based on satellite chlorophyll *a*, SST, bathymetry, and month, with mesh size set to 100 μm and all random effects set to zero (Table S2 in Supporting Information S1). Comparisons were used to assess the skill of the Historical 50-year time period from the ESM against the obsGLMM, both annually and seasonally. These skill statistics included a Taylor diagram, which is defined by Pearson's correlation coefficient, unbiased root mean square error, and normalized standard deviation, as well as Kendall's correlation coefficient, mean average error (MAE), and model bias. Supplemental skill assessments against obsMO and obsSM can be found in Appendix B.

2.6.2. Relationships Between Mesozooplankton and Chlorophyll

Finally, we assessed the multi-model ensemble to see if an emergent constraint could be derived to reduce uncertainty in future projections of mesozooplankton biomass. As mentioned previously, emergent constraints arise from: (a) physically interpretable patterns in the modeled historical simulations that are similar to those observed, and (b) a correlation between this historic pattern to a projected change under future conditions that is consistent across multiple models, even when individual patterns differ (Eyring et al., 2019; Hall & Qu, 2006; Hall et al., 2019; Kwiatkowski et al., 2017). For the observable historical relationship, we examined the relationship between mesozooplankton biomass and surface chlorophyll *a*, as recent modeling efforts have shown that it can be a promising constraint for model development (Luo et al., 2022). We contrasted the Historical 50-year mean (1965–2014) against the mean of the last 50 years of the SSP5-8.5 projections (2051–2100). The historic relationship was estimated with an ordinary least squares linear regression of $\log_{10}$-transformed mesozooplankton biomass as a function of $\log_{10}$-transformed chlorophyll *a* for the observational products and the historic ESM output, fit to the 50-year mean in each grid cell. The future change was calculated as the change

in mesozooplankton biomass under SSP5-8.5 from Historical using the 50-year means and $\log_{10}$-transforming before taking the difference. The climate sensitivity of the mesozooplankton biomass was calculated with an ordinary least squares linear regression of the future change in mesozooplankton biomass in each ESM against the historic regression slope coefficients of the mesozooplankton-chlorophyll relationship of each ESM. Analyses were performed globally and by biome.

Key to the application of the emergent constraints methodology for observationally uncertain quantities such as biological variables is to define uncertainty bounds around the observational constraints. This can be done using a combination of the internal uncertainty of the observations or observational products, or by using multiple sets of observational products, each with uncertainty bounds on them (e.g., Kwiatkowski et al., 2017). Unfortunately, the mixed model structure of the obsGLMM makes it difficult to add robust uncertainty bounds on the mesozooplankton-chlorophyll relationship, as the covariance structures for the random variables cannot be easily determined. Therefore, we added the obsMO and the obsSM to assess the range of historic estimates of the relationship between mesozooplankton and chlorophyll *a*. Similarly, these historic relationships were estimated with an ordinary least squares linear regression of the global climatology of mesozooplankton biomass as a function of surface chlorophyll. To be consistent with the empirical model (Strömberg et al., 2009), we used the SeaWiFS 1997–2010 mission climatology for the obsSM linear relationship and to define the biomes for partitioning the mesozooplankton regionally. The obsMO relationship between mesozooplankton biomass and surface chlorophyll *a* is technically two linear regressions, one (obsMO-M) with the MODIS chlorophyll that forced the obsGLMM (and associated biomes) and another (obsMO-S) with the SeaWiFS chlorophyll that forced the obsSM (and associated biomes).

We use chlorophyll *a* concentrations as an indicator of resource availability to mesozooplankton because: (a) it is more easily observed than phytoplankton biomass that depends on physical sampling, or primary production that relies on physical experiments or algorithms based on satellite chlorophyll that give disparate products (Saba et al., 2011) and (b) it was used to estimate mesozooplankton biomass in the obsGLMM of Heneghan et al. (2020) and the empirical model of Strömberg et al. (2009), allowing for consistency within our study. We recognize that there are caveats associated with using surface chlorophyll *a* as an index of food for mesozooplankton. For one, the carbon to chlorophyll ratio of phytoplankton varies as a function of organism size and the availability of light and nutrients (Sathyendranath et al., 2009). Additionally, there can be a variable number of trophic levels between phytoplankton and mesozooplankton. However, recent efforts have shown that chlorophyll may be a good predictor for fisheries production (Friedland et al., 2012; Park et al., 2019), suggesting that it may capture aspects of the system related to whether energy is directed up to higher trophic levels or is recycled around the microbial loop. Nonetheless, surface chlorophyll *a* concentrations are the best currently available observations spanning time and space.

3. Results

3.1. Historic Distributions of Mesozooplankton

The global mean simulated mesozooplankton biomass of historical annual climatologies (291–654 mgC m^{-2}) are within an order of magnitude of biomass estimated from the obsGLMM (534 mgC m^{-2}; Figures 1 and 2, Table S3 in Supporting Information S1), with the exception of CAN (41 mgC m^{-2}). ESM mesozooplankton reveal similar spatial patterns to each other and the obsGLMM of high biomass in temperate, upwelling, and shelf regions, and low biomass in the subtropical gyres (Figure 1). However, the range of mesozooplankton biomass from subtropical gyres to upwelling areas is greater in ESMs (1.06–2.69 orders of magnitude, mean 1.62, excluding CAN) compared to the obsGLMM (1.12 orders of magnitude).

3.2. Comparisons of Observed Versus Model Mesozooplankton

With the exception of CAN, all models perform reasonably well against the estimate of observed mesozooplankton biomass (Figures 3 and 4). CAN stands out in terms of its biases (Figure 3c), MAEs (Figures 3b and 4), and Pearson correlation (Figure 4). Excluding CAN, skill metrics spanned −0.09 to 0.55 (0.31) for (ranked) Kendall correlations (Figure 3a), 0.20–0.66 (0.34) for MAE (Figure 3b), −0.03 to −0.49 (−0.23) for bias (Figure 3c), −0.14 to 0.76 (0.29) for (linear) Pearson correlations, 0.22–0.77 (0.37) for RMSE, and 0.97–4.51 (1.7) for normalized standard deviation (Figure 4, Table S4 in Supporting Information S1). Though means across the annual and

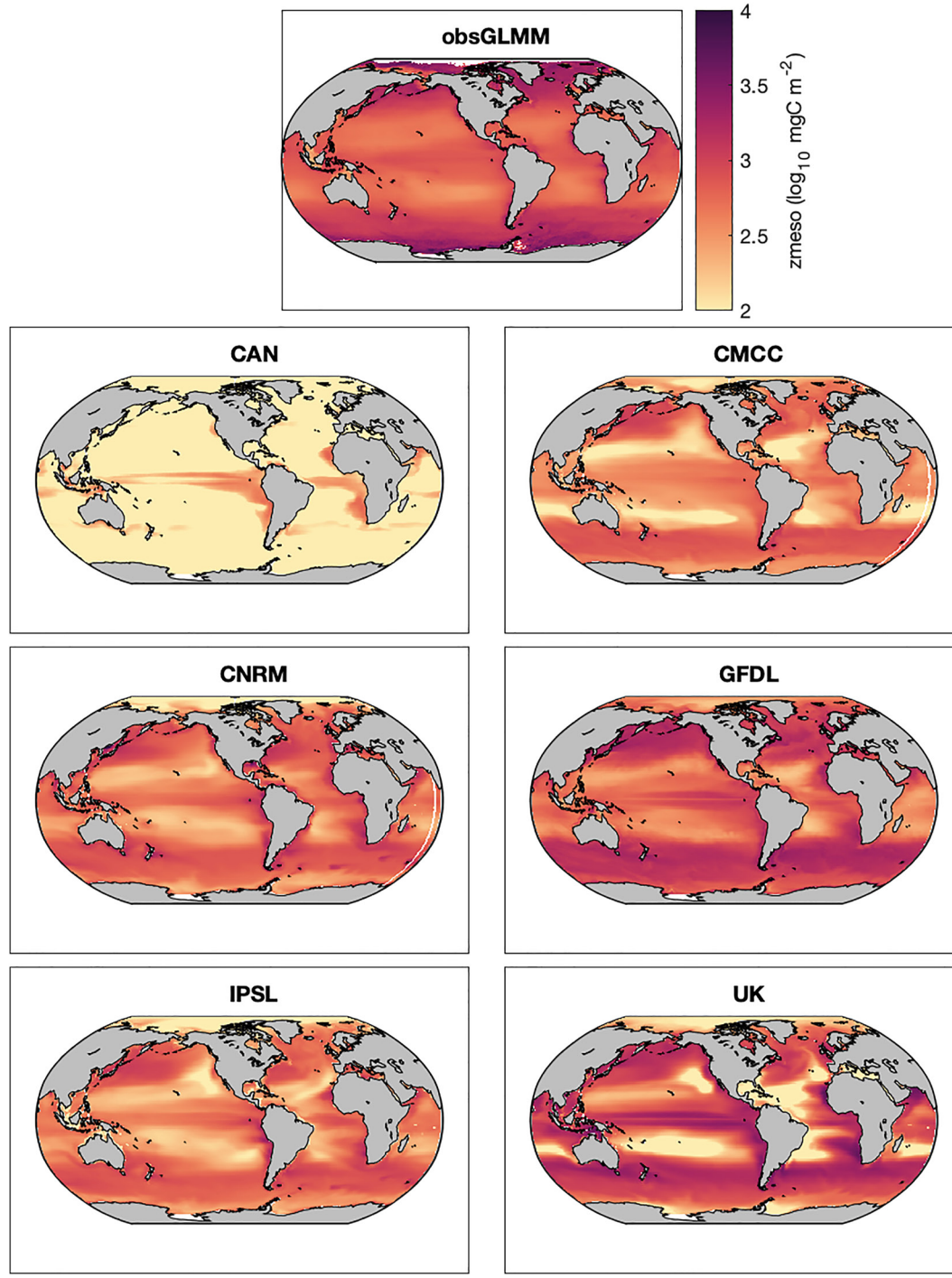

Figure 1. Distribution of mean mesozooplankton biomass ($\log_{10}$ mgC m^{-2}) in the obsGLMM and the ESMs over the Historical period (1965–2014).

seasonal climatologies are low, 11 out of 25 Pearson correlations are >0.3, showing moderate performance, 3 of which >0.5, indicating good performance (Figure 4, Table S4 in Supporting Information S1). Converting the untransformed MAE and bias to account for the log10-transformation equates to a 1.7–2.2x difference in biomasses simulated by the ESMs compared to the obsGLMM on average (Figures 3b and 3c). GFDL typically has the lowest mean absolute error and smallest bias, followed by CNRM, though the skill metric values and their rankings vary by season inconsistently across the ESMs (Figures 3b and 3c). The difference in the range of

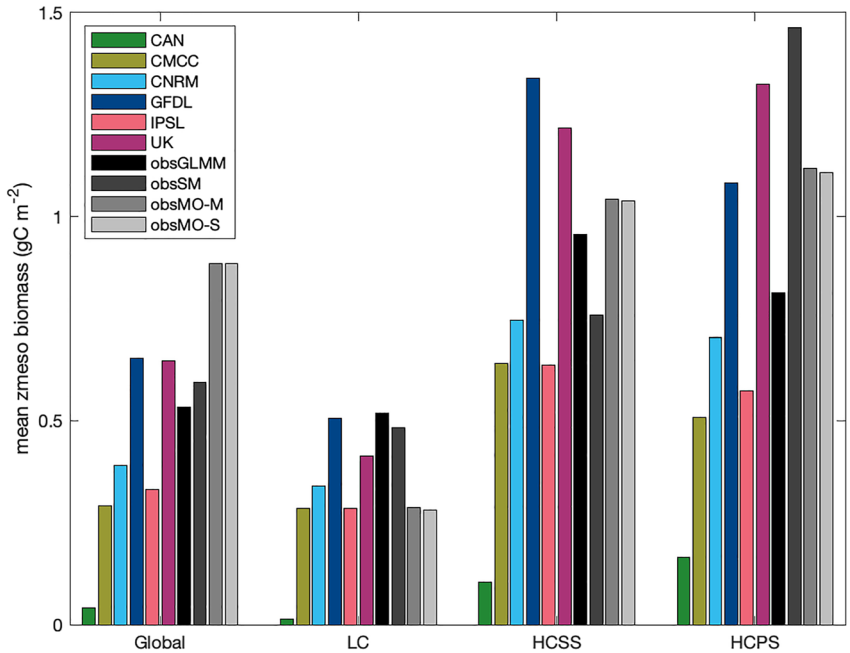

Figure 2. Area-weighted means of the full 12-month climatology of mesozooplankton biomass (mg C m^{-2}) during the 50-year Historical period (1965–2014), globally and by biome.

mesozooplankton biomass in the obsGLMM is noticeable in bias maps that show overestimates in temperate and upwelling regions of GFDL and UK and underestimates in the subtropics of all models (Figure S3 in Supporting Information S1). These differences in spatial variability are quantified with the normalized standard deviation, which are 1.7x greater on average across annual and seasonal climatologies (Figure 4). The large standard deviations stem from simulating a greater range in biomass than the obsGLMM, which is best illustrated with the higher highs and lower lows of UK (Figure 1) that tends to have the highest standard deviations and RMSE values (Figure 4). ESMs differ in the performance as measured by ranked and linear correlations. Pearson correlations give performance from best to worst as CMCC, GFDL, CNRM, IPSL, UK (Figure 4, Table S4 in Supporting Information S1), while Kendall correlations result in IPSL with the best skill (Figure 3a).

In terms of seasonal variability, some models perform better than others during certain seasons relative to the observation-based model (Figures 3 and 4). In general, ESMs have greater Pearson correlations with the obsGLMM during the meteorological fall (N hemisphere SON and S hemisphere MAM), followed by spring (Figure 4, Table S4 in Supporting Information S1). Though correlations tend to be low during winter, errors and biases are reduced (Figures 3b, 3c, and 4). The CMCC model has the two best seasonal estimates that appear closest to the reference on the Taylor diagram that summarizes correlation, error, and standard deviation together (Taylor, 2001), though the GFDL, CNRM, and IPSL models also clump together near the reference (Figure 4). CAN and UK are often outliers on the Taylor diagram due to their large errors and standard deviations (Figures 3b, 3c, and 4).

Examining the seasonal cycle of the different models by biome illustrates differences in temperature, chlorophyll *a*, and mesozooplankton biomass that may relate to the mesozooplankton skill. It is clear that CAN is an outlier in terms beyond mesozooplankton biomass (Figure 5a). CAN chlorophyll *a* is oddly higher than the other ESMs and the obsGLMM in the LC biome (Figure 5b, Figure S4 in Supporting Information S1). CAN SST is cooler than the others in the LC biome, but higher in both HC biomes (Figure 5c). The other five models and the obsGLMM are most similar in terms of SST. Similarities in SST across the models help unveil the relative influence of temperature versus resources on the mesozooplankton biomass since chlorophyll *a* varies more across the ESMs (Figure 5b). Seasonal cycles of chlorophyll *a* vary both in peak magnitude and timing (Figure 5b). The obsGLMM input chlorophyll *a* only shows a seasonal peak in the HCSS biome. The CNRM and IPSL models have less seasonal variability in chlorophyll *a* in all biomes and globally, while CMCC, GFDL, and UK demonstrate strong seasonal peaks in most biomes. Despite these differences, there are common seasonal progressions

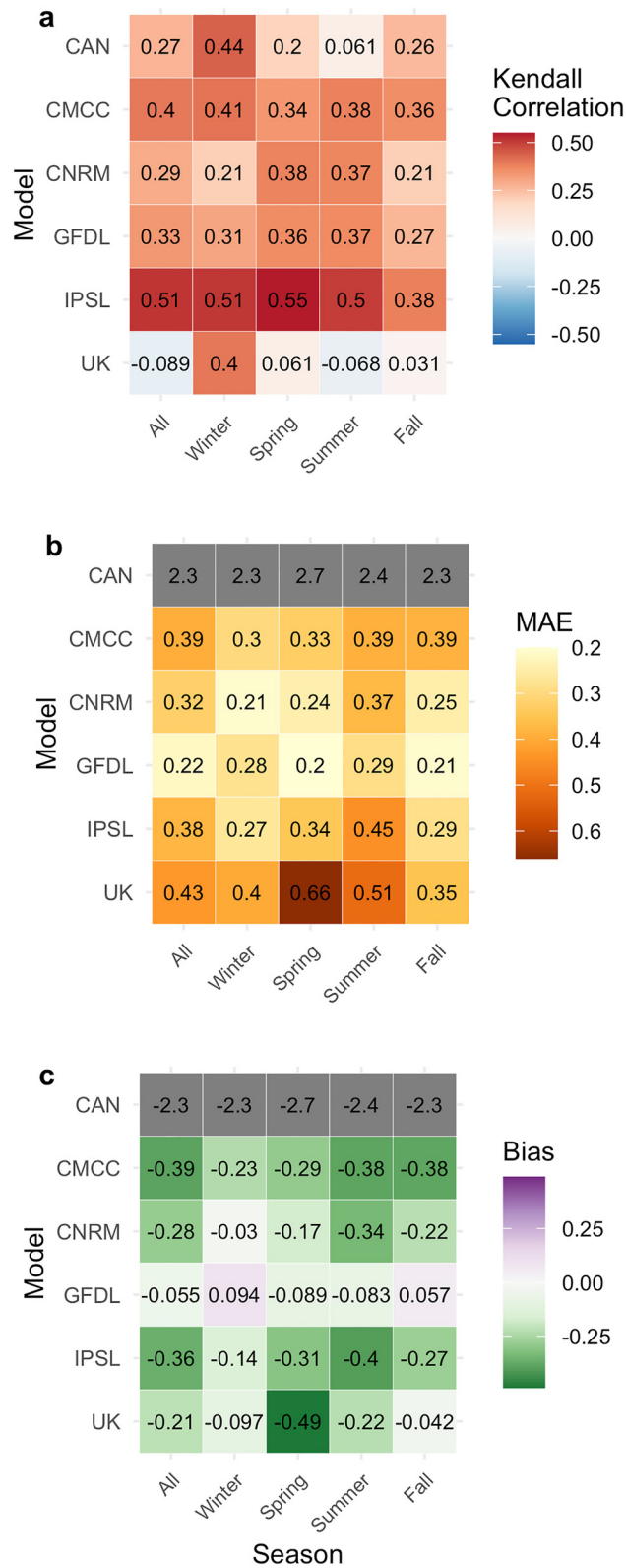

Figure 3. (a) Kendall correlation coefficient, (b) mean absolute error (MAE), and (c) bias of mesozooplankton in the Historical simulations compared to the obsGLMM. All biomasses were $\log_{10}$ transformed. Southern hemisphere seasonal climatologies were shifted 6 months (e.g., Win = N Hem DJF and S Hem JJA). Gray shading denotes values outside the range of the colorbar.

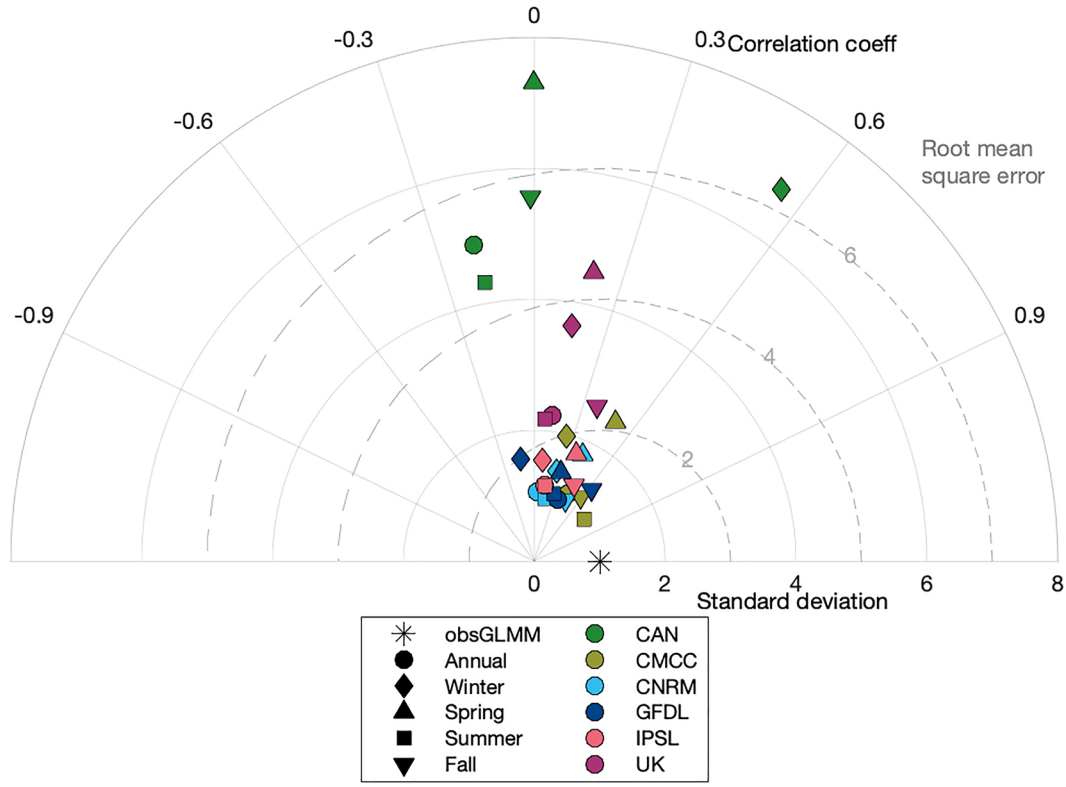

Figure 4. Taylor diagram skill of mesozooplankton in the Historical simulations compared to the obsGLMM. The angle gives correlation coefficients, the radial distance from the origin denotes normalized standard deviation (solid lines), and the radial distance from the observations is the RMSE (dashed lines). All biomasses were initially $\log_{10}$ transformed. Southern hemisphere seasonal climatologies were shifted 6 months (e.g., Win = N Hem DJF and S Hem JJA).

in the LC and HCSS biomes, with chlorophyll *a* peaking in the late winter/early spring in the LC biome, while in the HCSS biome it peaks in the late spring/early summer with a smaller bloom in the fall for some ESMs. This seasonal progression is mimicked in the mesozooplankton biomasses, with greater peaks in the HCSS biome (Figure 5a). In individual biomes, the range of mesozooplankton biomass varies less in CNRM, IPSL and the obsGLMM, like chlorophyll *a*, with the exception of the obsGLMM in HCSS. The CMCC mesozooplankton biomass also has less seasonal variability, but is in contrast to strong peaks in chlorophyll *a* in the HC biomes. GFDL and UK exhibit strong mesozooplankton seasonality and though they do not peak at the same time in each biome, the maxima are usually 1–2 months after the chlorophyll *a* bloom. In contrast, the peaks in chlorophyll *a* in the HCPS biome occur at different times of the year across the ESMs, yet the mesozooplankton biomass is smoother, more like SST. In most cases, the obsGLMM inputs of SST and chlorophyll *a* from MODIS-Aqua and estimates of mesozooplankton biomass are within the range of the ESMs and correspond to similar seasonal patterns. The earlier maxima in obsGLMM SST could be explained by poor satellite coverage in higher latitudes during winter months, which could lead to the earlier blooms of mesozooplankton globally and in the HCSS biome.

3.3. Projected Trends of Mesozooplankton and Its Drivers

The six ESMs with mesozooplankton respond differently under SSP5-8.5 future climate change simulations (Figure 6). With the exception of CAN, all models exhibit a decline in mesozooplankton biomass, but some models (CNRM, IPSL) decline much less than others (CMCC, GFDL, UK) (Figure 6a). The percent change in mesozooplankton biomass over time roughly mimics that of surface chlorophyll *a* (Figure 6b) for the CAN and IPSL models, but the relationships between mesozooplankton with chlorophyll *a* or temperature (Figure 6c) in the other ESMs are not straightforward. CNRM exhibits a decline in mesozooplankton biomass despite an increase in chlorophyll *a* and a smaller temperature change compared to the other ESMs. In contrast, the GFDL and UK

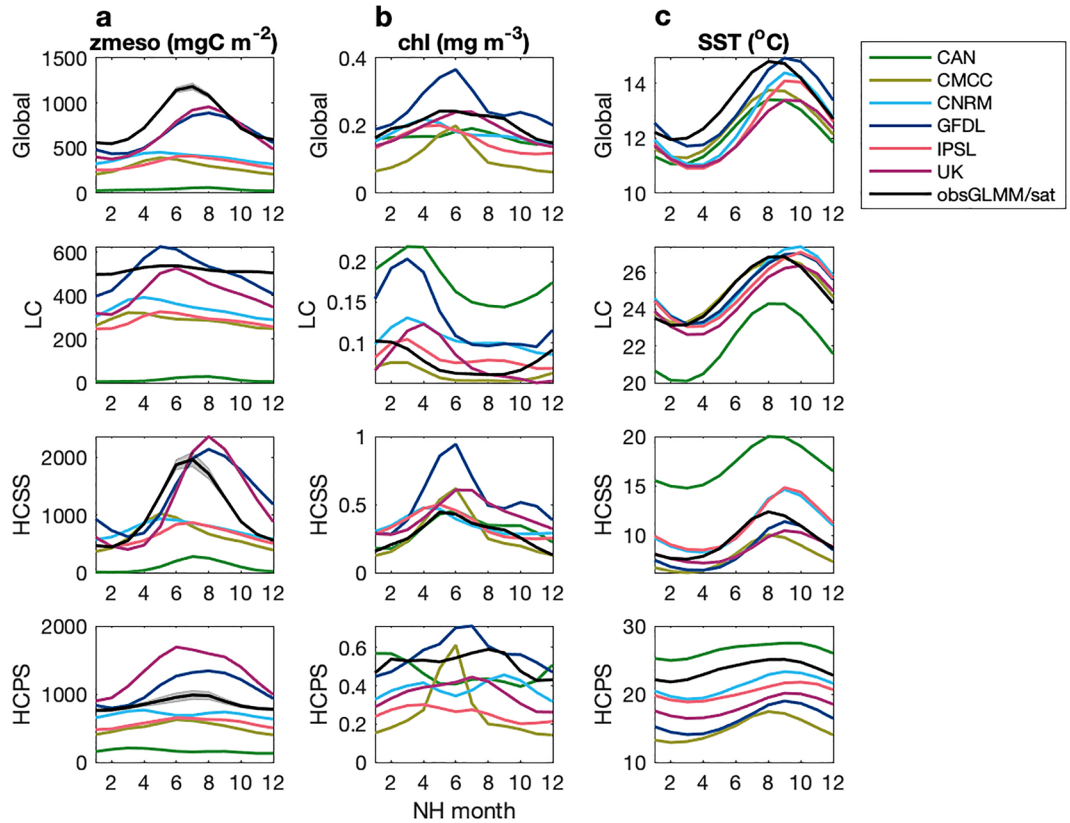

Figure 5. Annual climatologies of Historical (a) mesozooplankton biomass (mg C m^{-2}) in ESMs and the obsGLMM. (b) Surface chlorophyll a concentration (mg m^{-3}), and (c) sea surface temperature (SST) in ESMs and MODIS satellite observations (input to obsGLMM). Southern hemisphere seasonal climatologies were shifted 6 months (e.g., Win = N Hem DJF and S Hem JJA). LC regions above 45°N/below 45°S were excluded from the analysis. Shading around the obsGLMM is ±2 standard deviations.

models show large declines in mesozooplankton biomass, but are on opposite ends of the range for temperature change. Mesozooplankton simulated by GFDL decreases as much as or more than UK in spite of a smaller decline in chlorophyll *a* and a much weaker temperature response. And the greatest decrease in mesozooplankton occurs in the CMCC model, which has neither the greatest drop in chlorophyll *a* nor increase in temperature. The ESMs vary in their estimates of absolute amounts of mesozooplankton and chlorophyll *a*, both historically and in the future, but less so for temperature (Figures 6d–6f).

3.4. Relationships Between Mesozooplankton and Chlorophyll *a*

To better elucidate relationships between mesozooplankton and chlorophyll *a* for the ESMs, we excluded the CAN model from further analyses because of its poor model skill regarding the key variables in this study, although CAN-specific results can be found in Appendix A.

Using linear regression to quantify the historic relationships between mesozooplankton biomass and chlorophyll *a* concentration, we found that there are three models with slopes that fall between the obsGLMM and the obsSM relationships globally: CNRM, GFDL, and IPSL (Figures 7a and 7b). All ESMs had steeper relationships than the obsGLMM globally, as well as by biome (Table S5 in Supporting Information S1). The use of the obsSM allowed for an observational endmember with a slope greater than the obsGLMM that helped to constrain the models. The two obsMO-based constraints, the obsMO-S and obsMO-M, fell in the center between the obsGLMM and obsSM, with the three well constrained models clustering tightly around them. It should be noted that obsGLMM and obsSM are based on surface chlorophyll *a* and so it is unsurprising that the regressions are reasonably strong at a global scale (Figures S5f and S5g in Supporting Information S1) compared to obsMO (Figures S5h and S5i in Supporting Information S1).

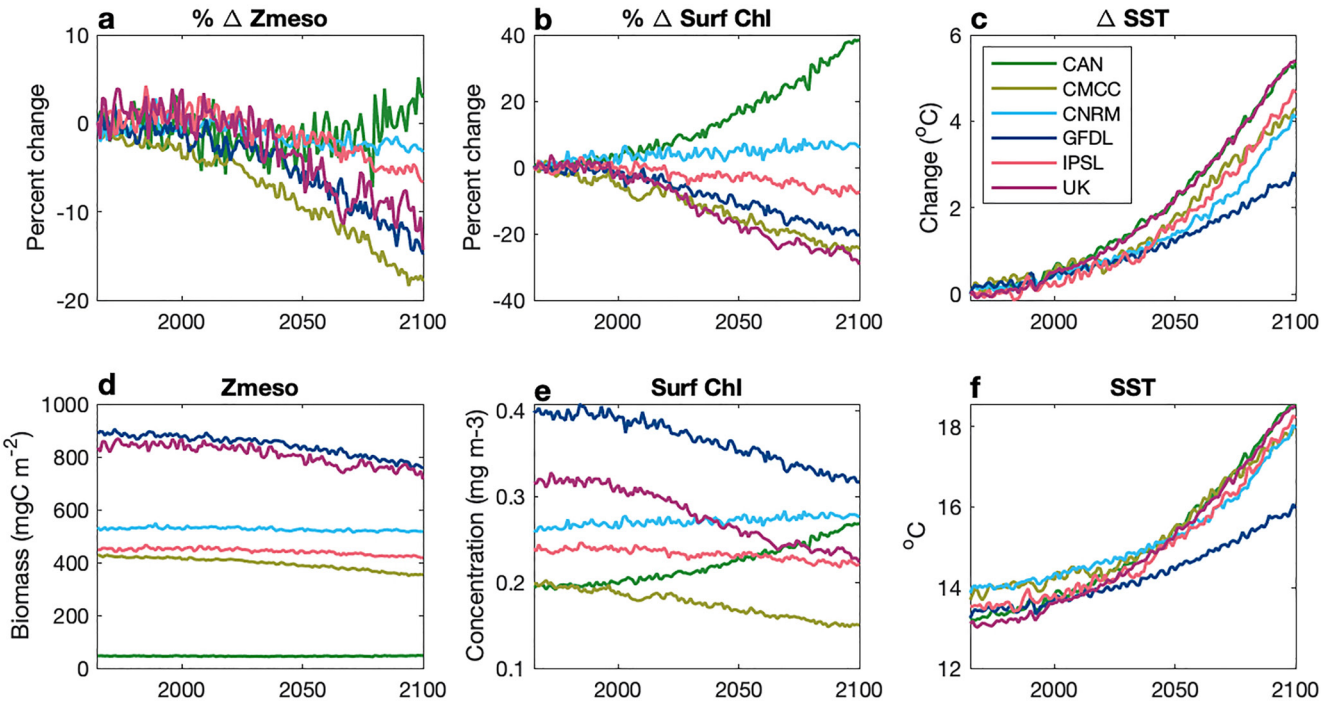

Figure 6. Globally averaged time series of ESM simulated (a and d) mesozooplankton biomass, (b and e) surface chlorophyll concentration, and (c and f) SST from 1965 to 2100. Shown are (a–c) percent change (zmeso, suf chl) and change (SST) from 1951 and (d–f) absolute values.

From the biome perspective, the tightest relationship between ESM mesozooplankton biomass and chlorophyll *a* is found in the LC biome (Figures S5a–S5e in Supporting Information S1). In this biome, the obsGLMM and obsSM have slopes of 0.41 and 0.80, respectively, with the GFDL model having the only slope (0.56; Table S5 in Supporting Information S1, Figures 7a and 7b) that fell within the bounds of the observational products. The other models exhibited slopes between 0.95 and 2.02 (Table S5 in Supporting Information S1). The scaling of the mesozooplankton-chlorophyll *a* relationship is weaker in the HCSS and HCPS biomes, becoming negative for CMCC and UK in the HCPS biome. In these biomes, the scaling relationships in the CNRM, GFDL, and IPSL models all fell within the observational bounds set by obsGLMM and obSM (Table S5 in Supporting Information S1). However, the scatter around these linear relationships denotes the greater variability at the regional level beyond biomes, which differs by ESM (Figures S5a–S5e in Supporting Information S1).

Finally, we compare the strength of the mesozooplankton-chlorophyll *a* relationship among the different models under historical conditions to the change in mesozooplankton biomass under future conditions. Though absolute values vary across models, there is a consistent pattern. The sensitivity of future projected mesozooplankton biomass on a global scale is related to the historical relationship between mesozooplankton and surface chlorophyll *a* ($r^2 = 0.49$, $p = 0.19$), where models that have a steeper mesozooplankton-chlorophyll *a* scaling slope exhibit greater declines in mesozooplankton biomass under climate change (Figure 7b). We found that the global climate sensitivity of mesozooplankton biomass is primarily driven by the LC biome, which exhibited greater changes in mesozooplankton biomass (Figure S7a in Supporting Information S1) and stronger mesozooplankton-chlorophyll *a* relationships (Table S5 in Supporting Information S1) than the HC biomes. This finding provides a potential framework for future targeted observations, as well as a potential emergent constraint by which future projections of mesozooplankton biomass change can be narrowed (Figures 7c and 7d).

4. Discussion

Zooplankton have not been considered previously in major comparisons of the skill of ESMs (Bopp et al., 2013; Fu et al., 2022; Kwiatkowski et al., 2020; Séférian et al., 2020). Here we assessed how well six ESMs from CMIP6 represent mesozooplankton under historical conditions, and found patterns across models, biomes, and seasons. Using an observational model (Heneghan et al., 2020), we show five out of six ESMs reproduced the

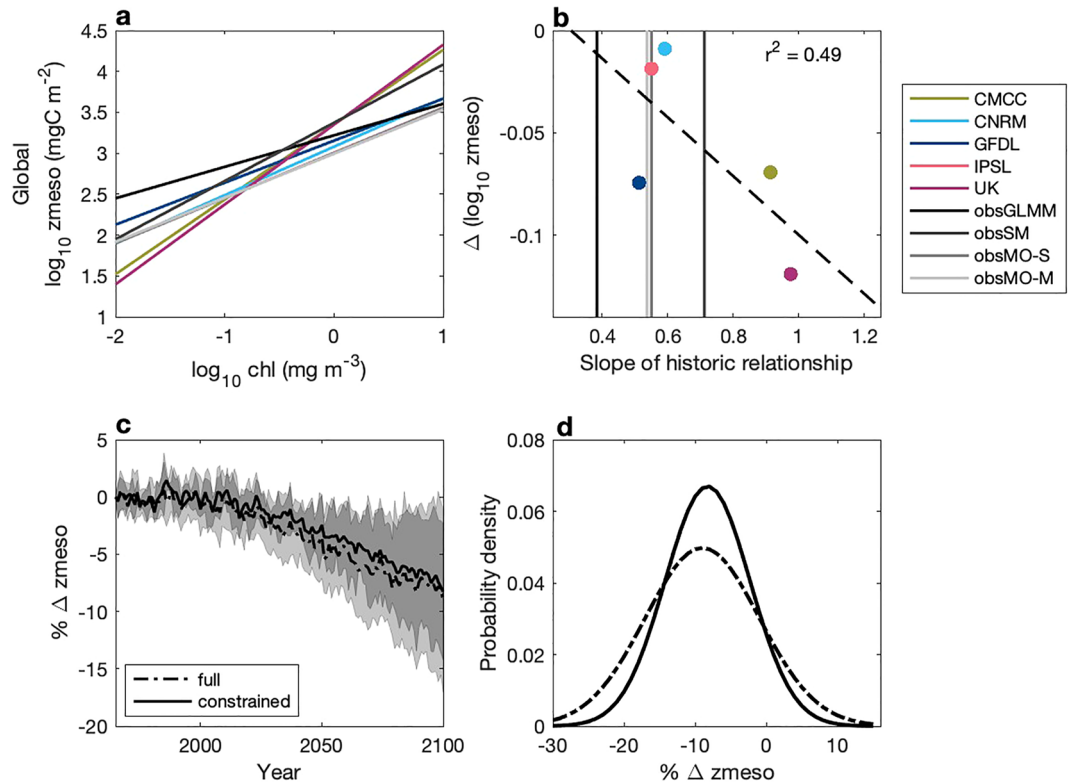

Figure 7. (a) Linear regressions of mesozooplankton biomass ($\log_{10}$ mg C m^{-2}) and surface chlorophyll a concentration ($\log_{10}$ mg m^{-3}) globally during the Historic period (1965–2014 mean). (b) Change in projected mesozooplankton biomass under SSP5.8-5 (2051–2100 mean) as a function of the relationship between mesozooplankton and chlorophyll under Historical conditions (linear regression slope). CAN was excluded as an outlier in (a and b). In (b), dashed black line: linear regression fit. obsGLMM: observation-based GLMM; obsSM: Strömberg et al. (2009) model; obsCS: COPEPOD data with SeaWiFS chlorophyll and related biomes; obsCM: COPEPOD data with MODIS chlorophyll and related biomes. Ensemble mean percent change in mesozooplankton biomass (c) from 1965 over time and (d) in the SSP5-8.5 (2051–2100) period compared to the Historical (1965–2014) period. In (c), the dashed line and light shading is the full, unconstrained mean ± 1 std of all six models, while the solid line and dark shading is the constrained mean ± 1 std of the CNRM, GFDL, and IPSL models. The means and standard deviations of (c) are represented as a probability density function in (d).

observed large-scale global pattern of mesozooplankton biomass with moderate skill (Figure 1). Further, all ESMs broadly reflected the positive scaling relationship between mesozooplankton biomass and surface chlorophyll *a* seen in observations under historical conditions (Figure 7a, Figures S5a–S5e and Table S5 in Supporting Information S1), though only three out of six ESMs exhibited scaling relationships within observational bounds set by the Heneghan et al. (2020) GLMM and the Strömberg et al. (2009) model (Figure 7b). This observational relationship, driven primarily by the low chlorophyll (LC) biome, allows for the application of an emergent constraint to reduce uncertainty in future projections of mesozooplankton biomass changes (Figures 7c and 7d).

4.1. Global Biomass Distributions in ESMs

Most ESMs produced reasonable estimates of mesozooplankton biomass that were within an order of magnitude of biomass estimated from the observational products and replicate large-scale patterns across biome. A key element in our model evaluation is the application of the biome methodology for assessing patterns in distinct parts of the ocean that are governed by fundamentally different dynamics. Spatially, all ESMs except CAN produced similar patterns of low biomass in subtropical gyres (e.g., LC biome) and high biomass in temperate areas (e.g., HCSS biome) and upwelling regions (e.g., HCPS biome). The biggest difference across models was the spatial range of mesozooplankton biomass, with for example, 2.68 orders of magnitude in UK but 1.06 orders of magnitude in GFDL between the 1st and 99th percentiles. In comparison, the obsGLMM exhibited 1.12 orders of magnitude in spatial variability, though the statistical model is likely dampening variability by smoothing over

the extremes. The individual observations from the COPEPOD database (Moriarty & O'Brien, 2013) span 2.38 orders of magnitude. Reproduction of the contrast in oligotrophic versus eutrophic regions is important for modeling upper trophic levels, particularly in productive ecosystems such as continental shelves. However, just because a model exhibits a larger dynamical range in the concentration of a quantity that better mimics the dynamical range of observations (and other skill metrics) does not necessarily mean that it is doing so by accurately representing the involved processes.

Mesozooplankton biomasses from all ESMs, except for the CAN model, were positively correlated with the observation-based statistical model (obsGLMM), as well as the other two observational products (Figure S6 in Supporting Information S1). Although Pearson correlations of annual means were relatively weak (<0.5), there were a few strongly correlated seasons and model errors and biases were often low. Model errors (RMSE, MAE, and bias) were greatest in summer and lowest in winter, whereas the highest Pearson correlations and greatest Taylor diagram skill were in the fall season. These skill estimates indicate that while the ESMs represented the annual mean state well, they exhibited biases in representing seasonal variations (spring bloom and winter die-off) relative to the GLMM observational product. However, the obsGLMM smooths over a large degree of spatial and temporal variability in observed zooplankton biomass, and thus may underestimate seasonal extremes. Nevertheless, our findings on how model skill differs seasonally could inform initialized Earth system model predictability and prediction experiments. These experiments generally investigate how predictable an Earth system quantity is within a certain time frame (e.g., subseasonal to interannual, interannual to decadal) and are ensembles of free-running simulations that are initialized from a forced model or observations (Yeager et al., 2018). Because of the high model skill in the fall, we recommend experiments targeting mesozooplankton (among other quantities) to be initialized at this time (e.g., Park et al., 2019; Yeager et al., 2018), as that season is one in which the models most closely match the observed ocean. This may have implications for evaluating fisheries predictions (Park et al., 2019; Stock et al., 2017).

4.2. Relationships Between Mesozooplankton and Chlorophyll

The relationship between mesozooplankton and chlorophyll *a* illustrates how mesozooplankton will respond given a certain change in chlorophyll *a*. We expected a positive relationship between spatial variations in mesozooplankton biomass and surface chlorophyll *a* (Richardson & Schoeman, 2004), as chlorophyll *a* can serve as a proxy for both phytoplankton biomass and primary production (Brewin et al., 2015; Friedland et al., 2012; Marañón et al., 2014), and the mesozooplankton-chlorophyll relationship has been used for recent model development efforts (Luo et al., 2022). While the positive relationship was generally true, particularly on the large scale, there were regional exceptions. Both CMCC and UK had negative relationships with chlorophyll *a* in the HCPS biome under historic conditions (Figures S5a, S5e, and Table S5 in Supporting Information S1).

There are several potential explanations for the variable and sometimes negative relationships between mesozooplankton biomass and chlorophyll *a* globally and in the biomes that can span many latitudes. For one, relationships between mesozooplankton and chlorophyll *a* will be affected by different parameterizations of the influence of temperature on physiological rates that may differ for phytoplankton and zooplankton and vary by model. Also, regional differences in light availability and ice cover, including differences in the ice dynamics of each ESM, may select for phytoplankton types with higher chlorophyll:carbon ratios (Sathyendranath et al., 2009). Thus, phytoplankton composition and the types of prey available to mesozooplankton will impact the mesozooplankton-chlorophyll relationship in all regions for each ESM. Both the effects of temperature and prey composition may be driving the two negative relationships in the HCPS biome. In these cases, it is likely that the inclusion of some arctic regions in the CMCC and UK HCPS biomes results in combining multiple mesozooplankton-chlorophyll *a* relationships into one (Figures S5a and S5e in Supporting Information S1). Temperature and prey composition could also account for the large degree in scatter within an ESM and differences in slopes across biomes and ESMs, as detailed below.

4.2.1. Role of Prey Composition

The strength of the relationships between mesozooplankton biomass and surface chlorophyll *a* and the potential influence of prey composition is exemplified by the sensitivity of mesozooplankton biomass to climate change (Figure 7b). The two models with reduced sensitivity, CNRM and IPSL, both use the PISCES2.0 BGC sub-model, and have small projected changes in chlorophyll *a* and even increased chlorophyll *a* in many oligotrophic regions

(Figure S7b in Supporting Information S1) or globally (CNRM, Figure 6b). These are regions where the size structure of phytoplankton should shift toward smaller phytoplankton that perform better under low nutrient, highly stratified conditions (Finkel et al., 2010). The ability of mesozooplankton to consume small phytoplankton separates the PISCES2.0 model from the BGC models of CAN, CMCC, and GFDL. This feature provides resilience against declines in large phytoplankton in CNRM and IPSL. Mesozooplankton in the UK model can also prey on small phytoplankton, but there is a stronger relationship with chlorophyll *a* compared to CNRM and IPSL (Figure 7a, Table S5 in Supporting Information S1), and chlorophyll *a* declines much more in projections (Figure 6b, Figure S7b in Supporting Information S1). Thus it is likely that both small and large phytoplankton have declined, therefore lessening the importance of small phytoplankton as a food source.

It is important to note that projected changes in chlorophyll *a* concentration may not necessarily denote a global increase or decrease in phytoplankton biomass, but a shift in phytoplankton composition. Diatoms/large phytoplankton have greater chlorophyll:carbon ratios than small phytoplankton (Geider et al., 1997; Sathyendranath et al., 2009), which is often parameterized in the ESMs (e.g., CMCC, CNRM, GFDL, and IPSL). As a result, a decrease in chlorophyll *a* could signify a shift from a diatom-dominated phytoplankton to a community dominated by small phytoplankton without any change in total phytoplankton carbon biomass. In such a situation, the same amount of carbon energy would need to pass through an additional trophic level (small phytoplankton to microzooplankton) to reach the mesozooplankton in models that do not allow for mesozooplankton grazing on small phytoplankton (CAN, CMCC, and GFDL), whereas in the converse case, small phytoplankton energy could pass directly to mesozooplankton (CNRM, IPSL, and UK). The more restrictive diets may account for the greater global declines of mesozooplankton biomass in future projections of CMCC and GFDL (Figures 6a and 7b), which is further illustrated by the greater percent changes in projected mesozooplankton biomass compared to chlorophyll (Figures S7a and S7b in Supporting Information S1). The effects of changes in phytoplankton composition on the mesozooplankton biomass via increases in trophic level would be further amplified by climate warming.

The strength of the mesozooplankton-chlorophyll relationships and the potential influence of prey composition are also evident in the seasonal climatologies (Figure 5) and long-term projections (Figure 6). On one end of the model spectrum, strong seasonal peaks of chlorophyll followed by large peaks in mesozooplankton simulated by GFDL and UK in the HC biomes suggest a bottom-up relationship of chlorophyll directly from large phytoplankton prey that have higher chlorophyll:carbon ratios and other photosynthesis parameters conducive to bloom formation (Geider et al., 1997). Conversely, the CNRM and IPSL models exhibit little seasonality in chlorophyll or mesozooplankton, which could be due to the parameters for small and large phytoplankton being too similar, leading to insufficient niche separation or overactive grazing, particularly by microzooplankton. The long-term decline in CNRM mesozooplankton biomass despite an increase in chlorophyll and a smaller warming compared to other ESMs also supports the hypothesis that their small phytoplankton are much more competitive relative to the large phytoplankton under future climate change, creating a system with more recycling and proportionally greater primary production being grazed by microzooplankton than mesozooplankton. Changes to the size structure of the plankton in these ESMs will be examined in a follow-up study to better elucidate the mechanisms affecting mesozooplankton biomass.

4.2.2. Role of Temperature Dependence

The temperature dependence of metabolic rates influences seasonal variation in primary and secondary production rates, how they respond to climate change, and the overall trophic transfer efficiency of the fixed carbon to higher trophic levels. The ESMs in this study all have slightly different temperature dependencies of phytoplankton growth, zooplankton grazing, and mortality (Table 1). While all models were fairly consistent in their representation of phytoplankton temperature dependence with some exceptions for mortality (Q_{10} between 1.88 and 2.0, but a greater range in the CAN Arrhenius equation), they differed substantially in their zooplankton temperature dependence. The GFDL and CMCC models used the same temperature dependence for phytoplankton and zooplankton, but CAN, CNRM/IPSL, and UK did not. In particular, UK does not include temperature dependence of zooplankton rate processes, while CAN only has temperature dependence on zooplankton respiration. Though the temperature dependence of zooplankton physiological rates can be highly variable depending on taxa (Ikeda, 1985; López-Urrutia et al., 2006), chemical processes are temperature dependent (Arrhenius, 1889), and there is strong empirical evidence for temperature dependent grazing rates (e.g., Hansen et al., 1997).

The role of temperature dependence can be illustrated by the UK model in terms of projected changes in SST, chlorophyll, and mesozooplankton biomass. The UK model had one of the greatest temperature increases under SSP5-8.5 compared to the other ESMs, and also the greatest percent declines in chlorophyll, with the fastest decline around the year 2050 (Figure 6b). However, percent declines in mesozooplankton biomass by 2100 in the UK model fell within the middle of the pack of the ESM responses, between the GFDL and IPSL models (Figure 6a). One of the major structural differences between the UK model and others is its lack of temperature dependence for zooplankton. No temperature dependence of zooplankton grazing means that as ocean temperatures warm, phytoplankton growth rates increase, but zooplankton grazing rates stay constant. Proportionally, zooplankton grazing pressure decreases, and that relative pressure decreases non-linearly with warming. Decreased grazing pressure can result in fast declines in phytoplankton populations, particularly in the nutrient-limited lower latitudes, as much of the microzooplankton grazing controls phytoplankton population growth by returning nutrients to the planktonic food web (Calbet, 2001; Calbet & Landry, 2004). This would be exemplified by an initial boom in phytoplankton biomass from the release of grazing, followed by a bust after using all the available nutrients. Accordingly, this can also result in faster expansions of the low chlorophyll biome (e.g., Figure S2 in Supporting Information S1). While it is not possible to completely isolate the impact of the lack of temperature dependence of zooplankton rates in a multimodel comparison, this mechanism may be one driver of the large differences between the UK ESM and other models, particularly with respect to the mesozooplankton-chlorophyll sensitivity.

4.3. Emergent Constraints and Reducing Model Ensemble Uncertainty

Emergent constraints are a relatively new tool for reducing the uncertainty of climate model ensemble projections (Eyring et al., 2019; Hall & Qu, 2006; Hall et al., 2019; Kwiatkowski et al., 2017). Such analyses isolate a relationship *within* individual ESMs that then correlate with a system specific response to climate change *across* ESMs. With an emergent constraint, ensemble members could then be removed from, or downweighted within, an ensemble if their historic relationship is not within the range of observations and/or their climate sensitivity is outside the uncertainty bounds of the emergent constraint. Here, we were able to identify a positive relationship between historic mesozooplankton biomass and surface chlorophyll *a* concentration in observations and two observation-based products that spanned productivity gradients across the global ocean. All six ESMs exhibited this positive scaling, though only three models (CNRM, GFDL, and IPSL) exhibited scaling relationships, as determined by the slope of the mesozooplankton-chlorophyll regression, that fell within the observational bounds. This relationship is admittedly weaker than other emergent relationships (Hall & Qu, 2006; Kwiatkowski et al., 2017; Terhaar et al., 2021), likely due to the small numbers of models with distinct mesozooplankton groups ($n = 5$, excluding CanESM) in the CMIP6 ensemble. Nonetheless, it provides an initial constraint on the future response of mesozooplankton to climate change such that ensembles of mesozooplankton biomass could be reduced to the CNRM, GFDL, and IPSL model outputs (Figures 7c and 7d). Still, reducing the ensemble to these three ESMs did not result in much difference between the unconstrained and constrained projected changes in mesozooplankton biomass under SSP5-8.5 (Figures 7c and 7d), as these three models still encompassed a large range of the climate response of mesozooplankton (Figure 6a). Since there were only a few ESMs with an explicit mesozooplankton group, this precluded the use of significance bounds of the emergent constraint to reduce model ensemble uncertainty beyond the observational constraints.

Across models, the strength of the scaling relationship was related to the change in mesozooplankton biomass under SSP5-8.5, with stronger relationships leading to larger declines (e.g., UK) and weaker relationships to lesser change (e.g., CNRM, IPSL). This global emergent constraint is strongly influenced by the oligotrophic, LC regions where chlorophyll *a* generally decreases under climate change (Figure S7b in Supporting Information S1), resulting in larger reductions of mesozooplankton in the models where mesozooplankton are more closely tied to chlorophyll *a* (Figure S7a in Supporting Information S1). This pattern is in accord with the processes of warming and stratification causing the greatest declines in large phytoplankton biomass and thus mesozooplankton biomass in the oligotrophic gyres. Intermediate changes in globally integrated mesozooplankton biomass in the future were seen in both the GFDL and CMCC models, which had weak and strong global scaling relationships respectively, thereby reducing the tightness of the emergent constraint correlation. As described in Section 4.2, there are numerous structural and parameter differences between these two models, and all six ESMs, that could account for their different mesozooplankton-chlorophyll a scalings and the future response of mesozooplankton biomass. The link between the mesozooplankton-chlorophyll relationship and the change in mesozooplankton biomass under climate change has the potential to serve as an important emergent constraint on climate change

projections of plankton dynamics. The utility of this emergent constraint will increase with more observations of mesozooplankton biomass (particularly if taken across a range of chlorophyll values), greater representation of mesozooplankton in ESMs, refinement of the mesozooplankton component in the six ESMs here, and isolating the mechanism(s) underpinning the mesozooplankton-chlorophyll relationship.

It is important to note that ESMs with closer agreement to any subset of global or regional observations will not necessarily have more skillful responses to climate projections (Stock et al., 2011). It is possible (though unlikely) that a model could misrepresent global phytoplankton and mesozooplankton distributions but still skillfully represent the mesozooplankton-chlorophyll scaling. It is therefore important that the emergent constraint be applied in conjunction with traditional phytoplankton skill assessments and the new mesozooplankton skill assessments described here. Furthermore, emergent constraints can be derived from pseudo-correlations, thus it is important to test the constraint in an independent model ensemble and to understand the mechanism behind the historic relationship (Terhaar et al., 2021). As CMIP6 was the first round of the intercomparison project to include a mesozooplankton output, there is no other ensemble of global models for testing this constraint.

4.4. Perspectives

4.4.1. Benchmarking Mesozooplankton in ESMs

The path to improving zooplankton in ESMs requires treating zooplankton as more than a closure term on phytoplankton by making more comparisons with data. Historically, the assessment of zooplankton in ESMs has been lacking (Everett et al., 2017). An examination of 153 biogeochemical models by Arhonditsis and Brett (2004) found that 95% made comparisons of model output with phytoplankton data, but <20% made comparisons to zooplankton data. Further, when zooplankton were assessed, they were more poorly simulated than almost any other state variable (Arhonditsis & Brett, 2004). A key reason behind this disparity is the lack of model-ready datasets for zooplankton, as satellite ocean color and derived estimates of NPP allow for much broader spatiotemporal assessment of phytoplankton biomass and productivity (Behrenfeld & Falkowski, 1997; Carr et al., 2006). However, as more zooplankton observations and standardized products become available (Buitenhuis et al., 2010; Heneghan et al., 2020; Moriarty & O'Brien, 2013; Strömberg et al., 2009), more assessment of zooplankton becomes possible. In addition to zooplankton biomass, model developers could also consider assessment of other terms, such as physiological rates (Heneghan et al., 2020; Hirst et al., 2003; Kiørboe & Hirst, 2014), their associated model parameters (Buitenhuis et al., 2006; Stock et al., 2014b; Rohr et al., 2022), and productivity ratios (Stock & Dunne, 2010).

Continued advancement of model development, assessment, and applications will require more observations and their synthesis, as well as coordinated investigation of common and disparate processes currently incorporated in ESMs representing mesozooplankton. Assessment of mesozooplankton biomass would benefit from greater long-term sampling at locations across different biomes. Further, observations need to be converted into model-ready products that standardize diverse sets of measurements and can be compared across models. Prior to model assessment, process studies are essential for estimating rates that are needed to constrain parameters during model development. As evidenced by the varied mechanisms that can influence the relationship between mesozooplankton biomass and chlorophyll concentration, increased measurement and estimation of growth rates, production rates, and respiration rates as well as their temperature-dependence are necessary to constrain models and isolate the processes involved. We therefore recommend continued collaboration between zooplankton modelers and observationalists for mutual improvement of both models and observations (e.g., Hjøllo et al., 2021) to better understand the processes structuring planktonic ecosystems (Everett et al., 2017). At the same time, we need to illuminate the mechanistic drivers of mesozooplankton responses to chlorophyll *a* and temperature across models. This could be through future multi-model intercomparison projects that provide zooplankton growth rates and loss terms, or by comprehensive experimental assessments like those conducted for global models of phytoplankton (Laufkötter et al., 2015) and higher trophic levels (Heneghan et al., 2021).

4.4.2. Recommendations for Users of ESM Mesozooplankton Output

Our work suggests that the use of simulated mesozooplankton biomass in ESMs is probably more suited to global applications rather than regional ones. ESMs are parameterized at the global scale to prioritize large-scale patterns in carbon and nutrient cycling, which can result in strong regional biases (Stock et al., 2011). Though spatial resolution is improving and many of the CMIP6 ESMs were run on native grids with <1° resolution, many

are still too coarse to adequately simulate coastal upwelling, coastal shelf circulation, eddies, and basin-shelf exchanges that are relevant for upper trophic levels (Drenkard et al., 2021; Stock et al., 2011). Additionally, ESMs do not accurately capture physical and biogeochemical exchanges at coastal boundaries that have large impacts on coastal fisheries production, such as riverine and sediment nutrient sources, neither in the historical nor future projections (e.g., Liu et al., 2021). Furthermore, studies using ESMs should rely on longer time series of simulations rather than short time periods of ≤10 years because out-of-phase climate oscillations could result in either over- or underestimation of climate change (Drenkard et al., 2021).

More specifically in terms of mesozooplankton, studies should use ESM ensembles that span the range of potential outcomes if possible (Drenkard et al., 2021; Stock et al., 2011) or select one ESM based on the research problem, but with ESM-specific caveats detailed (Kearney et al., 2021). The six CMIP6 ESMs examined in this study not only project various responses in the mesozooplankton, but also have unique combinations of high or low SST, chlorophyll, and mesozooplankton that could provide interesting contrasts or a robust ensemble mean. It is also important to consider the ecosystem processes under investigation and whether or not they are reflected in the BGC model structure (Kearney et al., 2021). For example, if an application of the mesozooplankton output is as fish prey under the assumption that it is representative of crustacean zooplankton such as copepods, then it may be better to only consider models that do not allow the mesozooplankton to consume small phytoplankton, as copepods cannot ingest picophytoplankton and inefficiently ingest much of the nanophytoplankton (Fuchs & Franks, 2010). On the other hand, use of an ESM where the mesozooplankton prey on small phytoplankton should acknowledge that these mesozooplankton are also representing gelatinous zooplankton such as appendicularians and tunicates that can feed on such small organisms (Conley et al., 2018). Even though most ESMs parameterize the mesozooplankton as crustaceans, they structurally represent all other consumers of phytoplankton and microzooplankton such as larval fish, chaetognaths, jellyfish, appendicularians, and tunicates. These various consumers differ in carbon content by >2 orders of magnitude (Heneghan et al., 2020), which could lead to under- or overestimates of the amount of energy available to higher trophic levels both now and in the future under climate change (Heneghan et al., 2021, 2022).

5. Conclusion

On the global scale, five of the six ESMs in CMIP6 that include mesozooplankton performed relatively well with respect to the observation-derived estimates of mesozooplankton biomass, in that biomasses were of the same order of magnitude, showed the same broad geographic patterns of highs and lows, and exhibited similar seasonality. The historic relationship between simulated mesozooplankton biomass and surface chlorophyll *a* concentration in three of the ESMs fell within the observed relationship globally and in high chlorophyll biomes. Furthermore, this relationship related to the change in mesozooplankton biomass under climate change, thereby potentially serving as an emergent constraint on climate change projections. This is a promising emergent constraint because the relationship between mesozooplankton biomass and surface chlorophyll *a* is observable with current sampling methods. The mechanistic underpinning of the mesozooplankton-chlorophyll relationship could not be determined in this analysis due to the variety of differences in structure and parameterizations between models. Note that there is even greater structural uncertainty in the biogeochemistry component of the entire CMIP6 ensemble as only the few ESMs with mesozooplankton were considered here. Model development and assessment will benefit from increased attention paid to prey preferences, food web structure, and zooplankton temperature sensitivity in addition to existing key parameters such as maximum ingestion rates and half-saturation constants. Expanded assessment of zooplankton in ESMs will improve the representation of phytoplankton to which they are intrinsically linked, thereby constraining and hopefully advancing estimates of carbon cycling.

Data Availability Statement

Earth system model outputs were downloaded from the ESGF-CoG (https://esgf-node.llnl.gov/search/cmip6/) and ISIMIP DKRZ (https://www.isimip.org/dashboard/accessing-isimip-data-dkrz-server/) servers. The model output DOIs are listed in Table S1 of Supporting Information S1. The obsGLMM is published on GitHub https://github.com/MathMarEcol/CMIP6_ZoopStatisticalModel/ (Everett et al., 2022). The Moriarty and O'Brien (2013) data set on zooplankton carbon biomass can be accessed at https://www.st.nmfs.noaa.gov/copepod/biomass/

biomass-fields.html. All other observations and model outputs are available at the websites cited in the text of their descriptions.

Acknowledgments
We thank CA Stock and T Rohr for helpful comments on earlier drafts, O Aumont, M Butenschön, J Christian, T Lovato, R Séférian, and A Yool for clarifications on their BGC sub-models, and two anonymous reviewers for their helpful critique. CMP was funded by NOAA Grants NA20OAR4310441 and NA20OAR4310442. JYL acknowledges support from the NOAA Marine Ecosystem Tipping Points Initiative. JDE was funded by Australian Research Council Discovery Project DP190102293. CSH acknowledges support from the Open Philanthropy and NSF Grant 2021686.

References

Allan, R. P., Hawkins, E., Bellouin, N., & Collins, B. (2021). IPCC, 2021: Summary for policymakers. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, et al. (Eds.), *Climate change 2021: The physical science basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (in press). Cambridge University Press.

Arhonditsis, G. B., & Brett, M. T. (2004). Evaluation of the current state of mechanistic aquatic biogeochemical modeling. *Marine Ecology Progress Series*, *271*, 13–26. https://doi.org/10.3354/meps271013

Arrhenius, S. (1889). On the reaction rate of the inversion of non-refined sugar upon souring. *Zeitschrift für Physikalische Chemie*, *4*(1), 226–248. https://doi.org/10.1515/zpch-1889-0416

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*, 2465–2513. https://doi.org/10.5194/gmd-8-2465-2015

Aumont, O., Maury, O., Lefort, S., & Bopp, L. (2018). Evaluating the potential impacts of the diurnal vertical migration by marine organisms on marine biogeochemistry. *Global Biogeochemical Cycles*, *32*(11), 1622–1643. https://doi.org/10.1029/2018GB005886

Banse, K. (1992). Grazing, temporal changes of phytoplankton concentrations, and the microbial loop in the open sea. In P. Falkowski & A. D. Woodhead (Eds.), *Primary productivity and biogeochemical cycles in the sea* (pp. 409–440). Plenum Press.

Behrenfeld, M. J., & Boss, E. S. (2018). Student's tutorial on bloom hypotheses in the context of phytoplankton annual cycles. *Global Change Biology*, *24*(1), 55–77. https://doi.org/10.1111/gcb.13858

Behrenfeld, M. J., & Falkowski, P. G. (1997). A consumer's guide to phytoplankton primary productivity models. *Limnology & Oceanography*, *42*(7), 1479–1491. https://doi.org/10.4319/lo.1997.42.7.1479

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J.-S. S. (2009). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology & Evolution*, *24*(3), 127–135. https://doi.org/10.1016/j.tree.2008.10.008

Bopp, L., Resplandy, L., Orr, J. C., Doney, S. C., Dunne, J. P., Gehlen, M., et al. (2013). Multiple stressors of ocean ecosystems in the 21st century: Projections with CMIP5 models. *Biogeosciences*, *10*, 6225–6245. https://doi.org/10.5194/bg-10-6225-2013

Brewin, R. J., Sathyendranath, S., Jackson, T., Barlow, R., Brotas, V., Airs, R., & Lamont, T. (2015). Influence of light in the mixed-layer on the parameters of a three-component model of phytoplankton size class. *Remote Sensing of Environment*, *168*, 437–450. https://doi.org/10.1016/j.rse.2015.07.004

Brunner, L., Pendergrass, A. G., Lehner, F., Merrifield, A. L., Lorenz, R., & Knutti, R. (2020). Reduced global warming from CMIP6 projections when weighting models by performance and independence. *Earth Syst. Dynam.*, *11*(4), 995–1012. https://doi.org/10.5194/esd-11-995-2020

Buitenhuis, E., Le Quéré, C., Aumont, O., Beaugrand, G., Bunker, A., Hirst, A., et al. (2006). Biogeochemical fluxes through mesozooplankton. *Global Biogeochemical Cycles*, *20*(2), GB2003. https://doi.org/10.1029/2005gb002511

Buitenhuis, E. T., Rivkin, R. B., Sailley, S., & Le Quéré, C. (2010). Biogeochemical fluxes through microzooplankton. *Global Biogeochemical Cycles*, *24*(4), GB4015. https://doi.org/10.1029/2009gb003601

Calbet, A. (2001). Mesozooplankton grazing effect on primary production: A global comparative analysis in marine ecosystems. *Limnology & Oceanography*, *46*(7), 1824–1830. https://doi.org/10.4319/lo.2001.46.7.1824

Calbet, A., & Landry, M. R. (2004). Phytoplankton growth, microzooplankton grazing, and carbon cycling in marine systems. *Limnology & Oceanography*, *49*(1), 51–57. https://doi.org/10.4319/lo.2004.49.1.0051

Carr, M.-E., Friedrichs, M. A., Schmeltz, M., Aita, M. N., Antoine, D., Arrigo, K. R., et al. (2006). A comparison of global estimates of marine primary production from ocean color. *Deep Sea Research Part II: Topical Studies in Oceanography*, *53*(5–7), 741–770. https://doi.org/10.1016/j.dsr2.2006.01.028

Christian, J. R., Denman, K. L., Hayashida, H., Holdsworth, A. M., Lee, W. G., Riche, O. G., et al. (2021). Ocean biogeochemistry in the Canadian Earth system model version 5.0. 3: CanESM5 and CanESM5-CanOE. *Geoscientific Model Development Discussions*, *15*(11), 4393–4424.

Conley, K. R., Lombard, F., & Sutherland, K. R. (2018). Mammoth grazers on the ocean's minuteness: A review of selective feeding using mucous meshes. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1878), 20180056. https://doi.org/10.1098/rspb.2018.0056

Drenkard, E. J., Stock, C., Ross, A. C., Dixon, K. W., Adcroft, A., Alexander, M., et al. (2021). Next-generation regional ocean projections for living marine resource management in a changing climate. *ICES Journal of Marine Science*, *78*(6), 1969–1987. https://doi.org/10.1093/icesjms/fsab100

Eppley, R. W. (1972). Temperature and phytoplankton growth in the sea. *Fisheries Bulletin*, *70*, 41063–41085.

Everett, J. D., Baird, M. E., Buchanan, P., Bulman, C., Davies, C., Downie, R., et al. (2017). Modeling what we sample and sampling what we model: Challenges for zooplankton model assessment. *Frontiers in Marine Science*, *4*, 77. https://doi.org/10.3389/fmars.2017.00077

Everett, J. D., Heneghan, R. F., & Richardson, A. J. (2022). The spatial distribution of zooplankton biomass (version 1) [Dataset]. GitHub. https://github.com/MathMarEcol/CMIP6_ZoopStatisticalModel

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Eyring, V., Cox, P. M., Flato, G. M., Gleckler, P. J., Abramowitz, G., Caldwell, P., et al. (2019). Taking climate model evaluation to the next level. *Nature Climate Change*, *9*(2), 102–110. https://doi.org/10.1038/s41558-018-0355-y

Finkel, Z. V., Beardall, J., Flynn, K. J., Quigg, A., Rees, T. A. V., & Raven, J. A. (2010). Phytoplankton in a changing world: Cell size and elemental stoichiometry. *Journal of Plankton Research*, *32*(1), 119–137. https://doi.org/10.1093/plankt/fbp098

Friedland, K. D., Stock, C., Drinkwater, K. F., Link, J. S., Leaf, R. T., Shank, B. V., et al. (2012). Pathways between primary production and fisheries yields of large marine ecosystems. *PLoS One*, *7*(1), e28945. https://doi.org/10.1371/journal.pone.0028945

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Olsen, A., et al. (2020). Global carbon budget 2020. *Earth System Science Data*, *12*(4), 3269–3340. https://doi.org/10.5194/essd-12-3269-2020

Frölicher, T. L., Rodgers, K. B., Stock, C. A., & Cheung, W. W. (2016). Sources of uncertainties in 21st century projections of potential ocean ecosystem stressors. *Global Biogeochemical Cycles*, *30*(8), 1224–1243. https://doi.org/10.1002/2015gb005338

Fu, W., Moore, J. K., Primeau, F., Collier, N., Ogunro, O. O., Hoffman, F. M., & Randerson, J. T. (2022). Evaluation of ocean biogeochemistry and carbon cycling in CMIP Earth system models with the International Ocean Model Benchmarking (IOMB) software system. *Journal of Geophysical Research: Oceans*, *127*(10), e2022JC018965. https://doi.org/10.1029/2022jc018965

Fu, W., Randerson, J., & Moore, J. (2015). Climate change impacts on net primary production (NPP) and export production (EP) regulated by increasing stratification and phytoplankton community structure in CMIP5 models. *Biogeosciences Discussions*, *12*, 12851–12897. https://doi.org/10.5194/bgd-12-12851-2015

Fuchs, H. L., & Franks, P. (2010). Plankton community properties determined by nutrients and size-selective feeding. *Marine Ecology Progress Series*, *413*, 1–15. https://doi.org/10.3354/meps08716

Geider, R. J., MacIntyre, H. L., & Kana, T. M. (1997). Dynamic model of phytoplankton growth and acclimation: Responses of the balanced growth rate and the chlorophyll a: Carbon ratio to light, nutrient-limitation and temperature. *Marine Ecology Progress Series*, *148*, 187–200. https://doi.org/10.3354/meps148187

Gentleman, W. C., & Neuheimer, A. B. (2008). Functional responses and ecosystem dynamics: How clearance rates explain the influence of satiation, food-limitation and acclimation. *Journal of Plankton Research*, *30*(11), 1215–1231. https://doi.org/10.1093/plankt/fbn078

Hall, A., Cox, P., Huntingford, C., & Klein, S. (2019). Progressing emergent constraints on future climate change. *Nature Climate Change*, *9*(4), 269–278. https://doi.org/10.1038/s41558-019-0436-6

Hall, A., & Qu, X. (2006). Using the current seasonal cycle to constrain snow albedo feedback in future climate change. *Geophysical Research Letters*, *33*(3), L03502. https://doi.org/10.1029/2005GL025127

Hansen, P. J., Bjørnsen, P. K., & Hansen, B. W. (1997). Zooplankton grazing and growth: Scaling within the 2–2,000-µm body size range. *Limnology & Oceanography*, *42*(4), 687–704. https://doi.org/10.4319/lo.1997.42.4.0687

Harrison, C. S., Luo, J. Y., Putman, N. F., Li, Q., Sheevam, P., Krumhardt, K., et al. (2021). Identifying global favourable habitat for early juvenile loggerhead sea turtles. *Journal of the Royal Society Interface*, *18*(175). https://doi.org/10.1098/rsif.2020.0799

Hatton, I. A., Heneghan, R. F., Bar-On, Y. M., & Galbraith, E. D. (2021). The global ocean size-spectrum from bacteria to whales. *bioRxiv*. https://doi.org/10.1101/2021.04.03.438320

Heneghan, R., Everett, J., Blanchard, J., Sykes, P., & Richardson, A. (2022). Climate-driven zooplankton shifts could cause global declines in food quality for fish. *ResearchSquare*, 0–25. (submitted). https://doi.org/10.21203/rs.3.rs-953268/v1

Heneghan, R. F., Everett, J. D., Sykes, P., Batten, S. D., Edwards, M., Takahashi, K., et al. (2020). A functional size-spectrum model of the global marine ecosystem that resolves zooplankton composition. *Ecological Modelling*, *435*, 109265. https://doi.org/10.1016/j.ecolmodel.2020.109265

Heneghan, R. F., Galbraith, E., Blanchard, J. L., Harrison, C., Barrier, N., Bulman, C., et al. (2021). Disentangling diverse responses to climate change among global marine ecosystem models. *Progress in Oceanography*, *198*, 102659. https://doi.org/10.1016/j.pocean.2021.102659

Henson, S., Laufkötter, C., Leung, S., Giering, S., Palevsky, H., & Cavan, E. (2021). What the flux? Uncertain response of ocean biological carbon export in a changing world. *Earth and Space Science Open Archive*. https://doi.org/10.1002/essoar.10507873.1

Hirst, A. G., Roff, J. C., & Lampitt, R. S. (2003). A synthesis of growth rates in marine epipelagic invertebrate zooplankton. *Advances in Marine Biology*, *44*, 1–142. https://doi.org/10.1016/s0065-2881(03)44002-9

Hjøllo, S. S., Hansen, C., & Skogen, M. D. (2021). Assessing the importance of zooplankton sampling patterns with an ecosystem model. *Marine Ecology Progress Series*, *680*, 163–176. https://doi.org/10.3354/meps13774

Ikeda, T. (1985). Metabolic rates of epipelagic marine zooplankton as a function of body mass and temperature. *Marine Biology*, *85*(1), 1–11. https://doi.org/10.1007/bf00396409

Kearney, K. A., Bograd, S. J., Drenkard, E., Gomez, F. A., Haltuch, M., Hermann, A. J., et al. (2021). Using global-scale Earth system models for regional fisheries applications. *Frontiers in Marine Science*, *1121*. https://doi.org/10.3389/fmars.2021.622206

Kiørboe, T., & Hirst, A. G. (2014). Shifts in mass scaling of respiration, feeding, and growth rates across life-form transitions in marine pelagic organisms. *The American Naturalist*, *183*(4), E118–E130. https://doi.org/10.1086/675241

Kwiatkowski, L., Aumont, O., & Bopp, L. (2019). Consistent trophic amplification of marine biomass declines under climate change. *Global Change Biology*, *25*(1), 218–229. https://doi.org/10.1111/gcb.14468

Kwiatkowski, L., Bopp, L., Aumont, O., Ciais, P., Cox, P. M., Laufkötter, C., et al. (2017). Emergent constraints on projections of declining primary production in the tropical oceans. *Nature Climate Change*, *7*(5), 355–358. https://doi.org/10.1038/nclimate3265

Kwiatkowski, L., Torres, O., Bopp, L., Aumont, O., Chamberlain, M., Christian, J., et al. (2020). Twenty-first century ocean warming, acidification, deoxygenation, and upper ocean nutrient decline from CMIP6 model projections. *Biogeosciences Discussions*, *17*(13), 3439–3470. https://doi.org/10.5194/bg-2020-16

Laufkötter, C., Vogt, M., Gruber, N., Aita-Noguchi, M., Aumont, O., Bopp, L., et al. (2015). Drivers and uncertainties of future global marine primary production in marine ecosystem models. *Biogeosciences*, *12*(23), 6955–6984. https://doi.org/10.5194/bg-12-6955-2015

Laufkötter, C., Vogt, M., Gruber, N., Aumont, O., Bopp, L., Doney, S. C., et al. (2016). Projected decreases in future marine export production: The role of the carbon flux through the upper ocean ecosystem. *Biogeosciences*, *13*, 19941–19998. https://doi.org/10.5194/bg-13-4023-2016

Liu, X., Stock, C. A., Dunne, J. P., Lee, M., Shevliakova, E., Malyshev, S., & Milly, P. C. D. (2021). Simulated global coastal ecosystem responses to a half-century increase in river nitrogen loads. *Geophysical Research Letters*, *48*(17), e2021GL094367. https://doi.org/10.1029/2021GL094367

Longhurst, A. (1995). Seasonal cycles of pelagic production and consumption. *Progress in Oceanography*, *36*(2), 77–167. https://doi.org/10.1016/0079-6611(95)00015-1

López-Urrutia, Á., San Martin, E., Harris, R. P., & Irigoien, X. (2006). Scaling the metabolic balance of the oceans. *Proceedings of the National Academy of Sciences*, *103*(23), 8739–8744. https://doi.org/10.1073/pnas.0601137103

Lotze, H. K., Tittensor, D. P., Bryndum-Buchholz, A., Eddy, T. D., Cheung, W. W., Galbraith, E. D., et al. (2019). Global ensemble projections reveal trophic amplification of ocean biomass declines with climate change. *Proceedings of the National Academy of Sciences*, *116*(26), 12907–12912. https://doi.org/10.1073/pnas.1900194116

Lovato, T., Peano, D., Butenschön, M., Materia, S., Iovino, D., Scoccimarro, E., et al. (2022). CMIP6 simulations with the CMCC Earth system model (CMCC-ESM2). *Journal of Advances in Modeling Earth Systems*, *14*(3), e2021MS002814. https://doi.org/10.1029/2021ms002814

Luo, J. Y., Stock, C. A., Henschke, N., Dunne, J. P., & O'Brien, T. D. (2022). Global ecological and biogeochemical impacts of pelagic tunicates. *Progress in Oceanography*, *205*, 102822. https://doi.org/10.1016/j.pocean.2022.102822

Maier-Reimer, E., & Hasselmann, K. (1987). Transport and storage of CO_2 in the ocean—An inorganic ocean-circulation carbon cycle model. *Climate Dynamics*, *2*(2), 63–90. https://doi.org/10.1007/bf01054491

Marañón, E., Cermeño, P., Huete-Ortega, M., López-Sandoval, D. C., Mouriño-Carballido, B., & Rodríguez-Ramos, T. (2014). Resource supply overrides temperature as a controlling factor of marine phytoplankton growth. *PLoS One*, *9*(6), e99312. https://doi.org/10.1371/journal.pone.0099312

Maury, O. (2010). An overview of APECOSM, a spatialized mass balanced "Apex Predators ECOSystem Model" to study physiologically structured tuna population dynamics in their ecosystem. *Progress in Oceanography*, *84*(1–2), 113–117. https://doi.org/10.1016/j.pocean.2009.09.013

Moriarty, R., & O'Brien, T. D. (2013). Distribution of mesozooplankton biomass in the global ocean. *Earth System Science Data*, *5*(1), 45–55. https://doi.org/10.5194/essd-5-45-2013

O'Brien, T. D. (2005). *COPEPOD: A global plankton database* (p. 136). US Department of Commerce, NOAA Technical Memoranda. O'Brien, T. D.: COPEPOD: The Global Plankton Data.
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
Petrik, C. M., Stock, C. A., Andersen, K. H., van Denderen, P. D., & Watson, J. R. (2019). Bottom-up drivers of global patterns of demersal, forage, and pelagic fishes. *Progress in Oceanography*, *176*, 102124. https://doi.org/10.1016/j.pocean.2019.102124
Postel, L., Fock, H., & Hagen, W. (2000). Biomass and abundance. In R. Harris, P. Wiebe, J. Lenz, H. R. Skjoldal, & M. Huntley (Eds.), *ICES zooplankton methodology manual* (pp. 83–192). Academic Press.
Richardson, A. J., & Schoeman, D. S. (2004). Climate impact on plankton ecosystems in the Northeast Atlantic. *Science*, *305*(5690), 1609–1612. https://doi.org/10.1126/science.1100958
Rohr, T., Richardson, A. J., Lenton, A., Shadwick, E., & Chamberlain, M. (2022). Recommendations for the formulation of grazing in marine biogeochemical and ecosystem models. *Progress in Oceanography*, *208*, 102878. https://doi.org/10.1016/j.pocean.2022.102878
Ryther, J. H. (1969). Photosynthesis and fish production in the sea. *Science*, *166*(3901), 72–76. https://doi.org/10.1126/science.166.3901.72
Saba, V. S., Friedrichs, M. A., Antoine, D., Armstrong, R. A., Asanuma, I., Behrenfeld, M. J., et al. (2011). An evaluation of ocean color model estimates of marine primary productivity in coastal and pelagic regions across the globe. *Biogeosciences*, *8*(2), 489–503. https://doi.org/10.5194/bg-8-489-2011
Sarmiento, J. L., Orr, J. C., & Siegenthaler, U. (1992). A perturbation simulation of CO_2 uptake in an ocean general circulation model. *Journal of Geophysical Research*, *97*(C3), 3621–3645. https://doi.org/10.1029/91jc02849
Sarmiento, J. L., Slater, R., Barber, R., Bopp, L., Doney, S. C., Hirst, A. C., et al. (2004). Response of ocean ecosystems to climate warming. *Global Biogeochemical Cycles*, *18*(3), GB3003. https://doi.org/10.1029/2003gb002134
Sathyendranath, S., Stuart, V., Nair, A., Oka, K., Nakane, T., Bouman, H., et al. (2009). Carbon-to-chlorophyll ratio and growth rate of phytoplankton in the sea. *Marine Ecology Progress Series*, *383*, 73–84. https://doi.org/10.3354/meps07998
Séférian, R., Berthet, S., Yool, A., Palmiéri, J., Bopp, L., Tagliabue, A., et al. (2020). Tracking improvement in simulated marine biogeochemistry between CMIP5 and CMIP6. *Current Climate Change Reports*, *6*(3), 95–119. https://doi.org/10.1007/s40641-020-00160-0
Sieburth, J. M., Smetacek, V., & Lenz, J. (1978). Pelagic ecosystem structure: Heterotrophic compartments of the plankton and their relationship to plankton size fractions. *Limnology & Oceanography*, *23*(6), 1256–1263. https://doi.org/10.4319/lo.1978.23.6.1256
Steinberg, D. K., & Landry, M. R. (2017). Zooplankton and the ocean carbon cycle. *Annual Review of Marine Science*, *9*(1), 413–444. https://doi.org/10.1146/annurev-marine-010814-015924
Stock, C., & Dunne, J. (2010). Controls on the ratio of mesozooplankton production to primary production in marine ecosystems. *Deep Sea Research Part I: Oceanographic Research Papers*, *57*(1), 95–112. https://doi.org/10.1016/j.dsr.2009.10.006
Stock, C. A., Alexander, M. A., Bond, N. A., Brander, K. M., Cheung, W. W., Curchitser, E. N., et al. (2011). On the use of IPCC-class models to assess the impact of climate on living marine resources. *Progress in Oceanography*, *88*(1–4), 1–27. https://doi.org/10.1016/j.pocean.2010.09.001
Stock, C. A., Dunne, J. P., Fan, S., Ginoux, P., John, J., Krasting, J. P., et al. (2020). Ocean biogeochemistry in GFDL's Earth System Model 4.1 and its response to increasing atmospheric CO_2. *Journal of Advances in Modeling Earth Systems*, *12*(10). https://doi.org/10.1029/2019MS002043
Stock, C. A., Dunne, J. P., & John, J. G. (2014a). Drivers of trophic amplification of ocean productivity trends in a changing climate. *Biogeosciences*, *11*(24), 7125–7135. https://doi.org/10.5194/bg-11-7125-2014
Stock, C. A., Dunne, J. P., & John, J. G. (2014b). Global-scale carbon and energy flows through the marine planktonic food web: An analysis with a coupled physical-biological model. *Progress in Oceanography*, *120*, 1–28. https://doi.org/10.1016/j.pocean.2013.07.001
Stock, C. A., John, J. G., Rykaczewski, R. R., Asch, R. G., Cheung, W. W., Dunne, J. P., et al. (2017). Reconciling fisheries catch and ocean productivity. *Proceedings of the National Academy of Sciences*, *114*(8), E1441–E1449. https://doi.org/10.1073/pnas.1610238114
Strömberg, K. H. P., Smyth, T. J., Allen, J. I., Pitois, S., & O'Brien, T. D. (2009). Estimation of global zooplankton biomass from satellite ocean colour. *Journal of Marine Systems*, *78*(1), 18–27. https://doi.org/10.1016/j.jmarsys.2009.02.004
Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, *106*(D7), 7183–7192. https://doi.org/10.1029/2000jd900719
Terhaar, J., Frölicher, T. L., & Joos, F. (2021). Southern Ocean anthropogenic carbon sink constrained by sea surface salinity. *Science Advances*, *7*(18), eabd5964. https://doi.org/10.1126/sciadv.abd5964
Tittensor, D. P., Eddy, T. D., Lotze, H. K., Galbraith, E. D., Cheung, W., Barange, M., et al. (2018). A protocol for the intercomparison of marine fishery and ecosystem models: Fish-MIP v1. 0. *Geoscientific Model Development*, *11*(4), 1421–1442. https://doi.org/10.5194/gmd-11-1421-2018
Tittensor, D. P., Novaglio, C., Harrison, C. S., Heneghan, R. F., Barrier, N., Bianchi, D., et al. (2021). Next-generation ensemble projections reveal higher climate risks for marine ecosystems. *Nature Climate Change*, *11*(11), 1–9. https://doi.org/10.1038/s41558-021-01173-9
Wiebe, P. H., & Benfield, M. C. (2003). From the Hensen net toward four-dimensional biological oceanography. *Progress in Oceanography*, *56*(1), 7–136. https://doi.org/10.1016/s0079-6611(02)00140-4
Yeager, S. G., Danabasoglu, G., Rosenbloom, N. A., Strand, W., Bates, S. C., Meehl, G. A., et al. (2018). Predicting near-term changes in the Earth system: A large ensemble of initialized decadal prediction simulations using the community Earth system model. *Bulletin of the American Meteorological Society*, *99*(9), 1867–1886. https://doi.org/10.1175/bams-d-17-0098.1
Yool, A., Palmiéri, J., Jones, C. G., de Mora, L., Kuhlbrodt, T., Popova, E. E., et al. (2021). Evaluating the physical and biogeochemical state of the global ocean component of UKESM1 in CMIP6 historical simulations. *Geoscientific Model Development*, *14*(6), 3437–3472.
Yool, A., Popova, E. E., & Anderson, T. R. (2011). Medusa-1.0: A new intermediate complexity plankton ecosystem model for the global domain. *Geoscientific Model Development*, *4*(2), 381–417. https://doi.org/10.5194/gmd-4-381-2011
Yool, A., Popova, E. E., & Anderson, T. R. (2013). MEDUSA-2.0: An intermediate complexity biogeochemical model of the marine carbon cycle for climate change and ocean acidification studies. *Geoscientific Model Development*, *6*(5), 1767–1811. https://doi.org/10.5194/gmd-6-1767-2013

References From the Supporting Information

Boucher, O., Denvil, S., Caubel, A., & Foujols, M. A. (2018). *IPSL IPSL-CM6A-LR model output prepared for CMIP6 CMIP historical. Version 20201026*. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.5195
Boucher, O., Denvil, S., Levavasseur, G., Cozic, A., Caubel, A., Foujols, M.-A., et al. (2019). *IPSL IPSL-CM6A-LR model output prepared for CMIP6 ScenarioMIP ssp585. Version 20210118*. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.5271

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., et al. (2020). Presentation and evaluation of the IPSL-CM6A-LR climate. *Journal of Advances in Earth Modeling Systems*, *12*(7). https://doi.org/10.1029/2019MS002010

Cherchi, A., Fogli, P. G., Lovato, T., Peano, D., Iovino, D., Gualdi, S., et al. (2019). Global mean climate and main patterns of variability in the CMCC-CM2 coupled model. *Journal of Advances in Modeling Earth Systems*, *11*(1), 185–209. https://doi.org/10.1029/2018MS001369

Good, P., Sellar, A., Tang, Y., Rumbold, S., Ellis, R., Kelley, D., & Kuhlbrodt, T. (2019). *MOHC UKESM1.0-LL model output prepared for CMIP6 ScenarioMIP ssp585. Version 20201120*. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.6405

Held, I. M., Guo, H., Adcroft, A., Dunne, J. P., Horowitz, L. W., Krasting, J. P., et al. (2019). Structure and performance of GFDL's CM4.0 climate model. *Journal of Advances in Modeling Earth Systems*, *11*(11), 3691–3727. https://doi.org/10.1029/2019MS001829

John, J. G., Blanton, C., McHugh, C., Radhakrishnan, A., Rand, K., Vahlenkamp, H., et al. (2018). *NOAA-GFDL GFDL-ESM4 model output prepared for CMIP6 ScenarioMIP ssp585. Version 20200621*. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.8706

Krasting, J. P., John, J. G., Blanton, C., McHugh, C., Nikonov, S., Radhakrishnan, A., et al. (2018). *NOAA-GFDL GFDL-ESM4 model output prepared for CMIP6 CMIP historical. Version 20200606*. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.8597

Lovato, T., Peano, D., & Butenschön, M. (2021a). *CMCC CMCC-ESM2 model output prepared for CMIP6 CMIP historical. Version 20200715*. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.13195

Lovato, T., Peano, D., & Butenschön, M. (2021b). *CMCC CMCC-ESM2 model output prepared for CMIP6 ScenarioMIP ssp585. Version 20200715*. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.13259

Séférian, R. (2018). *CNRM-CERFACS CNRM-ESM2-1 model output prepared for CMIP6 CMIP historical. Version 20201016*. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.4068

Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., Colin, J., et al. (2019). Evaluation of CNRM Earth-system model, CNRM-ESM2-1: Role of Earth system processes in present-day and future climate. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4182–4227. https://doi.org/10.1029/2019MS001791

Sellar, A. A., Jones, C. G., Mulcahy, J., Tang, Y., Yool, A., Wiltshire, A., et al. (2019). UKESM1: Description and evaluation of the UK Earth system model. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4513–4558. https://doi.org/10.1029/2019MS0017392019

Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., et al. (2019a). The Canadian Earth system model version 5 (CanESM5.0.3). *Geoscientific Model Development*, *12*(11), 4823–4873. https://doi.org/10.5194/gmd-12-4823-2019

Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., et al. (2019b). *CCCma CanESM5-CanOE model output prepared for CMIP6 CMIP historical. Version 20201016*. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.1026

Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., et al. (2019c). *CCCma CanESM5-CanOE model output prepared for CMIP6 ScenarioMIP ssp585. Version 20201120*. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.10276

Tang, Y., Rumbold, S., Ellis, R., Kelley, D., Mulcahy, J., Sellar, A., et al. (2019). *MOHC UKESM1.0-LL model output prepared for CMIP6 CMIP historical. Version 20201016*. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.6113

Voldoire, A. (2019). *CNRM-CERFACS CNRM-ESM2-1 model output prepared for CMIP6 ScenarioMIP ssp585. Version 20201120*. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.4226