

Assessment and constraint of mesozooplankton in CMIP6 Earth system models

C.M. Petrik^{1,2}, J.Y. Luo³, R.F. Heneghan⁴, J.D. Everett^{5,6,7}, C.S. Harrison⁸, A.J. Richardson^{5,6}

¹Department of Oceanography, Texas A&M University, College Station, TX, USA

²Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA

³NOAA Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA

⁴School of Mathematical Sciences, Queensland University of Technology, Brisbane, QLD, Australia

⁵School of Mathematics and Physics, The University of Queensland, St. Lucia, QLD, Australia

⁶Commonwealth Scientific and Industrial Research Organisation (CSIRO) Oceans and Atmosphere, Queensland Biosciences Precinct, St Lucia, QLD, Australia

⁷Centre for Marine Science and Innovation, The University of New South Wales, Sydney, NSW, Australia

⁸Department of Ocean and Coastal Science, Center for Computation and Technology, Louisiana State University, Baton Rouge, LA, USA

corresponding author: Colleen M. Petrik (cpetrik@ucsd.edu)

ORCID

C.M. Petrik – 0000-0003-3253-0455

R.F. Heneghan – 0000-0001-7626-1248

J.Y. Luo – 0000-0002-0032-9370

J.D. Everett – 0000-0002-6681-8054

C.S. Harrison – 0000-0003-4544-947X

A.J. Richardson – 0000-0002-9289-7366

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 vs. 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

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 microscopic 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 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 (Frolicher et al., 2016; Kwiatkowski et al., 2020; Tittensor et al., 2021; Fu et al., *in revision*). Multimodel ensembles are often used to capture a wider range of the different types of uncertainty. However, multimodel ensembles could be biased by unskillful models, which has led to the use of emergent constraints to reduce projection uncertainty by weighting models based on performance (Kwiatkowski et al., 2017; Eyring et al., 2019; Hall et al., 2019). We use this approach here, estimating performance by evaluating the mesozooplankton output from individual ESMs against observations through skill assessment and in comparison with other ESMs through emergent constraints analysis.

In this study we assess mesozooplankton in ESMs to answer four key questions. (1) How well do ESMs represent mesozooplankton under historic conditions? (2) Are there regions or seasons where ESM representation of mesozooplankton is better or worse? (3) Are ESMs able to reproduce the relationships between mesozooplankton biomass and surface chlorophyll *a* seen in observations? (4) 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 6th 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). 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 (Allen 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), surface chlorophyll concentration (chlos), sea surface temperature (tos), depth level thickness (thkcello), and mixed layer depth (mlotst).

The ocean biogeochemistry (BGC) sub-models of each ESM are described to varying degrees in model description publications (Table 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 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.

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
BFMv5.2 (CMCC)	2 NanoP, Diatoms	2 MicroZ, MesoZ
CanOE (CAN)	2 SmallP, LargeP	2 SmallZ, LargeZ
COBALTv2 (GFDL)	3 Diazotrophs, SmallP, LargeP	3 SmallZ (<200 m), MediumZ* (200-2000 m)
MEDUSA2.1 (UK)	2 Non-Diatoms, Diatoms	2 MicroZ, MesoZ

PISCES2.0 (CNRM, IPSL)	2	2
	NanoP, Diatoms	MicroZ, MesoZ

Note. g_{max} : maximum grazing rate; I_{max} : maximum ingestion rate; k : half saturation constant for grazing or ingestion rate. *Mesozooplankton in COBAL combines medium and large zooplankton. (a) k = Potential specific growth rate / Specific search volume, where Specific search volume = $0.025 \text{ m}^3 \text{ mgC}^{-1} \text{ d}^{-1}$, (b) $k = -\ln(0.5)/a$, where $a = 0.021 \text{ m}^3 \text{ mgC}^{-1}$ converted from $0.25 \text{ (mmol C m}^{-3})^{-1}$, (c) Converted from $1.25 \text{ mol N kg}^{-1}$ assuming Redfield Ratio of 106:16:1 (C:N:P), (d) Converted from $0.3 \text{ mmol N m}^{-3}$ assuming Redfield Ratio of 106:16:1 (C:N:P), (e) Converted from $20 \text{ } \mu\text{mol C L}^{-1}$. All conversions to mgC used $\text{mg} = \mu\text{mol} * 12.0107 * 1\text{e-}3$ or $\text{mg} = \text{nmol} * 12.0107 * 1\text{e-}6$.

In addition to varied diets, ingestion by zooplankton is modelled with different functional responses. BFMv5.2, CanOE, and COBALv2 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, COBALv2, and MEDUSA2.1 use some form of prey preference. The prey preference in COBAL 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 (COBALv2, MEDUSA2.1). Though models differ in which ratios they consider (C:N:Fe – CanOE; C:N:P – BFMv5.2, COBALv2; 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 COBALv2 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 COBALv2 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 Observation-based statistical model

We compared the historical mesozooplankton biomass from ESMs with an observation-based statistical model. We used the Generalized Linear Mixed Model (GLMM) of Heneghan et al. (2020), which produces global maps of zooplankton biomass for particular months or seasons, and is based on point observations of zooplankton biomass from the COPEPOD database (Moriarty & O’Brien, 2012). COPEPOD is a compilation of >196,000 individual observations of biomass from over 70 years. Rather than use the raw zooplankton biomass observations for model assessment, Heneghan et al. (2020) developed a statistical modelling approach based on the observations for several reasons. First, using raw observations is challenged by the fact that zooplankton observations are collected with >50 different sampling methods, including optical instruments and nets with different mesh sizes, mouth sizes, tow speeds, and tow directions (Wiebe & Benfield, 2003). Second, samples are collected at different times of day, seasons, and ocean depths, all of which should be considered as they influence zooplankton biomass. Last, zooplankton biomass has been measured in various ways, such as settled volume, displacement volume, wet weight, dry weight, ash free dry weight, carbon, and biovolume (Postel et al., 2000). Conversions between these methods of measuring biomass are based on small sample sizes from particular regions (Wiebe, 1988; Postel, 1990; Bode et al., 1998) and often provide negative biomass estimates when applied outside their region. There are simply no conversions to standardize for all these sampling differences and the diverse zooplankton communities present in distinct regions.

GLMMs provide a robust approach that adjusts for observation biases (Bolker et al., 2009). We describe here the GLMM of Heneghan et al. (2020) (hereafter called obsGLMM). The obsGLMM estimates observed zooplankton biomass measurements (the response) from the COPEPOD database as a function of environmental variables whilst adjusting for sampling biases. Here we assume that the obsGLMM estimates mesozooplankton biomass because most observations in COPEPOD are for mesozooplankton, with few microzooplankton captured because the meshes are too coarse. However, some macrozooplankton will be sampled (Wiebe & Benfield, 2003), while on the other hand mesozooplankton

could be underestimated because of net avoidance (Sameoto et al., 2000). The obsGLMM adjusts for temporal, spatial, and sampling biases within the dataset by including a suite of fixed and random predictors. 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), the mean depth of sampling (continuous), sampling day of year (continuous), and time of day (continuous) to adjust for diel vertical migration. To account for the circular nature of time of day (24-hr cycle), it was transformed using a truncated Fourier series. 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.

Environmental predictors in the obsGLMM included sea surface temperature, surface chlorophyll *a*, and bathymetry. The obsGLMM was fit using monthly climatologies of sea surface temperature and satellite chlorophyll *a* measurements 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. 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 is a seasonal climatology.

ObsGLMM captures most of the variability in the COPEPOD data, with an R^2 of 91% (82% from fixed effects). Yet, there are caveats associated with the obsGLMM. For one, the obsGLMM from Heneghan et al. (2020) is an estimate of mean biomass (based on 196,907 observations), 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 patchy distribution of zooplankton observations in time and space, the multiple methods of measuring biomass, and the different biases such as net mesh size associated with different observations (Everett et al., 2017). 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 Processing for model-observation comparisons

ObsGLMM 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 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 to a common $1^\circ \times 1^\circ$ grid. Similarly, total mesozooplankton biomass in the top 200 m from the GLMM was calculated on the same $1^\circ \times 1^\circ$ grid. Few COPEPOD data were collected before 1950 and the last addition of new observations was in 2015. For comparison with the obs-GLMM estimates, the 50-year period from 1965-2014 was analyzed from the Historical ESM simulations. We computed seasonal and annual climatologies from all 600 months of ESM output, standardizing southern hemisphere seasons to align with the northern hemisphere (e.g., winter corresponds to northern hemisphere December-February and southern hemisphere June-August). Absolute biomasses were transformed with a fourth-root before comparison to create a more normal frequency distribution to apply Pearson correlation and other statistics that assume normality. As absolute values between ESMs and observations might be different, but general patterns of relatively high and low biomass might be similar, we also scaled the ESM and obsGLMM outputs between -1 and 1.

2.4 Spatial scale and biome definition

The coarse spatial resolution of global ESMs ($\sim 25\text{-}100$ 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): (1) LC – low chlorophyll; (2) HCPS – high chlorophyll, permanently stratified; and (3) 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 $0.125 \text{ mg chl m}^{-3}$, 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, 1994; 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 (S  f  rian et al. 2020, Fu et al. *in review*), 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 $0.125 \text{ mg chl m}^{-3}$ 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-yr 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 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 vs. 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-yr 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 S1). However, there were still 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 labelled as High Chlorophyll, Permanently Stratified (HCPS) in some models. Still, the choice of using mixed layer depth to define permanently vs. 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.

2.5 Analyses

2.5.1 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 m and all random effects set to zero (Table S2). Comparisons were used to assess the skill of the Historical 50-yr time period from the ESM against the obsGLMM, both annually and seasonally. These skill statistics included Pearson’s correlation coefficient, unbiased root mean square error, normalized standard deviation (summarized in Taylor diagrams), and model bias.

2.5.2 Relationships between mesozooplankton and chlorophyll

Finally, we assessed if an emergent constraint exists in the relationship between mesozooplankton biomass and surface chlorophyll *a*. Such emergent constraints are named for their ability to provide additional constraints on projections, thereby reducing their uncertainty (Kwiatkowski et al., 2017; Eyring et al., 2019; Hall et al., 2019). Constraints emerge from: (1) patterns in historical simulations that are similar to those observed, and (2) a link between this historic pattern to a projected change under future conditions that is consistent across multiple models, even when individual patterns differ. We contrasted the Historical 50-yr 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 a 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-yr 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-yr means and \log_{10} -transforming before taking the difference. The climate sensitivity of the mesozooplankton biomass was calculated with a 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 a second empirical model of mesozooplankton biomass (obsSM), which is an observationally-based ($n=4,843$) trophic transfer model forced with SeaWiFS satellite ocean color data (Strömberg et al., 2009), to assess the range of historic estimates of the relationship between mesozooplankton and chlorophyll *a*. Similarly, this historic relationship was estimated with a linear regression of the global climatology of mesozooplankton biomass from the Strömberg et al. (2009) product as a function of surface chlorophyll using the SeaWiFS 1997-2010 mission climatology. The SeaWiFS climatology was also used to define the biomes for partitioning the Strömberg et al. (2009) mesozooplankton regionally. We additionally include a third historic relationship from the COPEPOD carbon biomass (mgC m^{-3} , $n=10,117$) dataset of mesozooplankton abundance (Moriarty & O'Brien, 2012) integrated over the top 200 m. This is technically two relationships, one with the MODIS chlorophyll that forced the obsGLMM (and associated biomes) and another 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: (1) 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 (2) 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 historical annual climatologies of simulated mesozooplankton biomass (291-654 mgC m⁻²) are within an order of magnitude of absolute biomass estimated from the obsGLMM (534 mgC m⁻²; Figure 1, Table 2), with the exception of CAN (41 mgC m⁻²). ESM zooplankton 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 (Figure 2a,c,e, Figure 3A). CAN stands out in terms of its biases and error (Figs 2c,e). GFDL has the smallest annual bias, followed by UK, CNRM, IPSL, and CMCC (Figure 2e). These biases and their rankings vary by season, with the exception of the GFDL model that always has the smallest bias (Figure 2e). The difference in the range of 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). These differences in spatial variability are quantified with the normalized standard deviation, which are greater than 1 for many ESMs and their seasonal climatologies (Figure 3a).

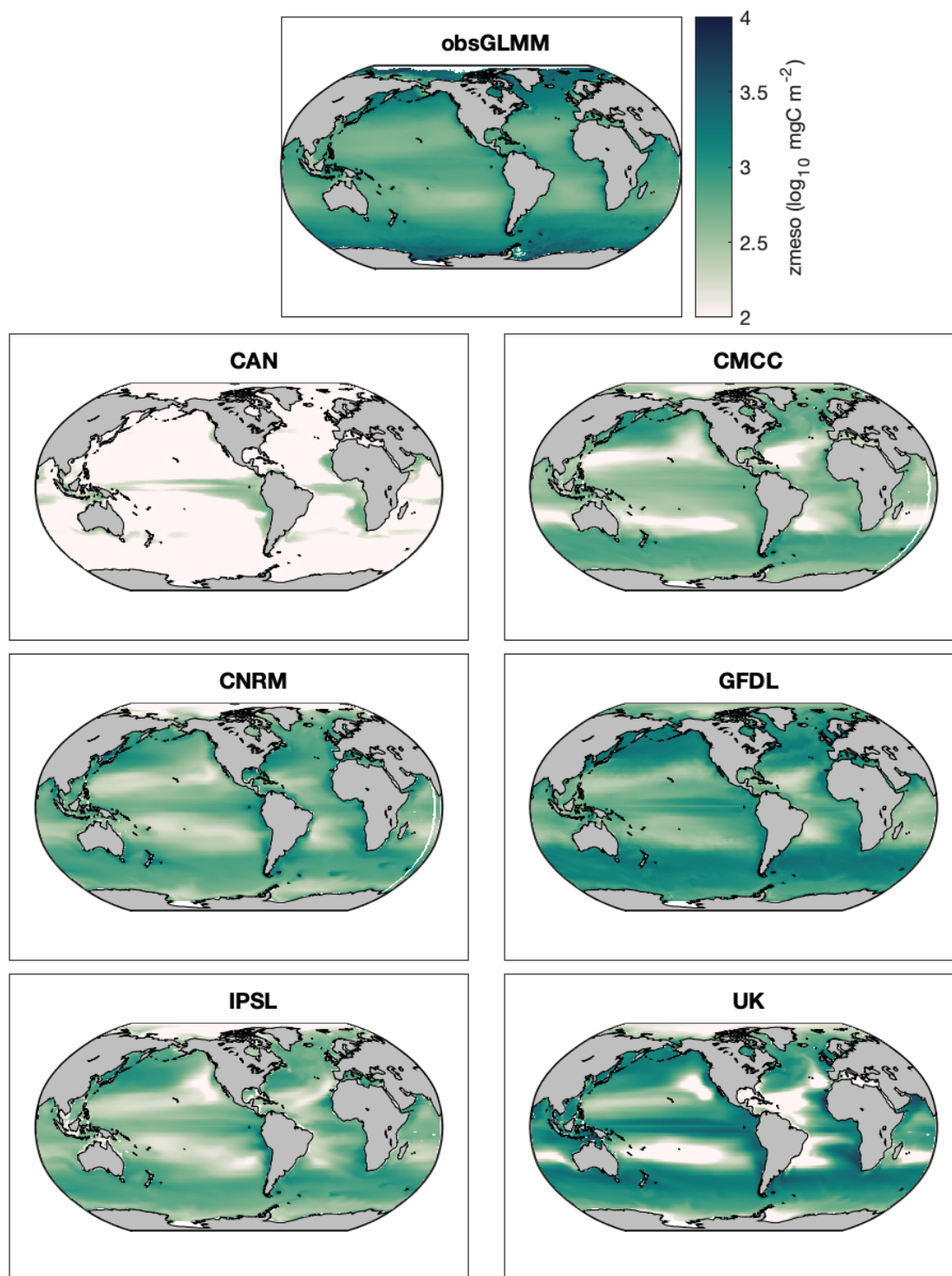


Figure 1. Distribution of mean mesozooplankton biomass (\log_{10} mgC m⁻²) over the Historical period (1965-2014) and the obsGLMM.

Table 2. *Historic area-weighted mean mesozooplankton biomass (mg C m⁻²) globally and by biome.*

Model/Obs	Global	LC	HCSS	HCPS
CAN	41	14	104	165
CMCC	291	285	641	509
CNRM	390	340	746	703
GFDL	654	507	1340	1083
IPSL	331	285	637	574
UK	646	414	1217	1325
obsGLMM	534	518	956	813
obsSM	595	483	758	1463
obsCM	884	287	1042	1118
obsCS	884	281	1039	1107

Note. CAN: CanESM5-CanOE; CMCC: CMCC-ESM2; CNRM: CNRM-ESM2-1; GFDL: GFDL-ESM4; IPSL: IPSL-CM6A-LR; UK: UKESM1-0-LL; obsGLMM: observation-based statistical model of Heneghan et al. (2020); obsSM: observation-based empirical model of Strömberg et al. (2009); obsCS: COPEPOD data with SeaWiFS chlorophyll and related biomes; obsCM: COPEPOD data with MODIS chlorophyll and related biomes.

In terms of seasonal variability, some models perform better than others during certain seasons relative to the observation-based model (Figure 2a,c,e, Figure 3a). In general, ESMs have greater correlations with the obsGLMM during the meteorological fall (N hemisphere SON and S hemisphere MAM), followed by spring (Figure 2a, Figure 3a). Though correlations tend to be low during winter, errors and biases are reduced (Figure 2c,e). When comparing absolute biomasses, the CMCC model generally has the lowest root mean square errors (Figure 2c) and

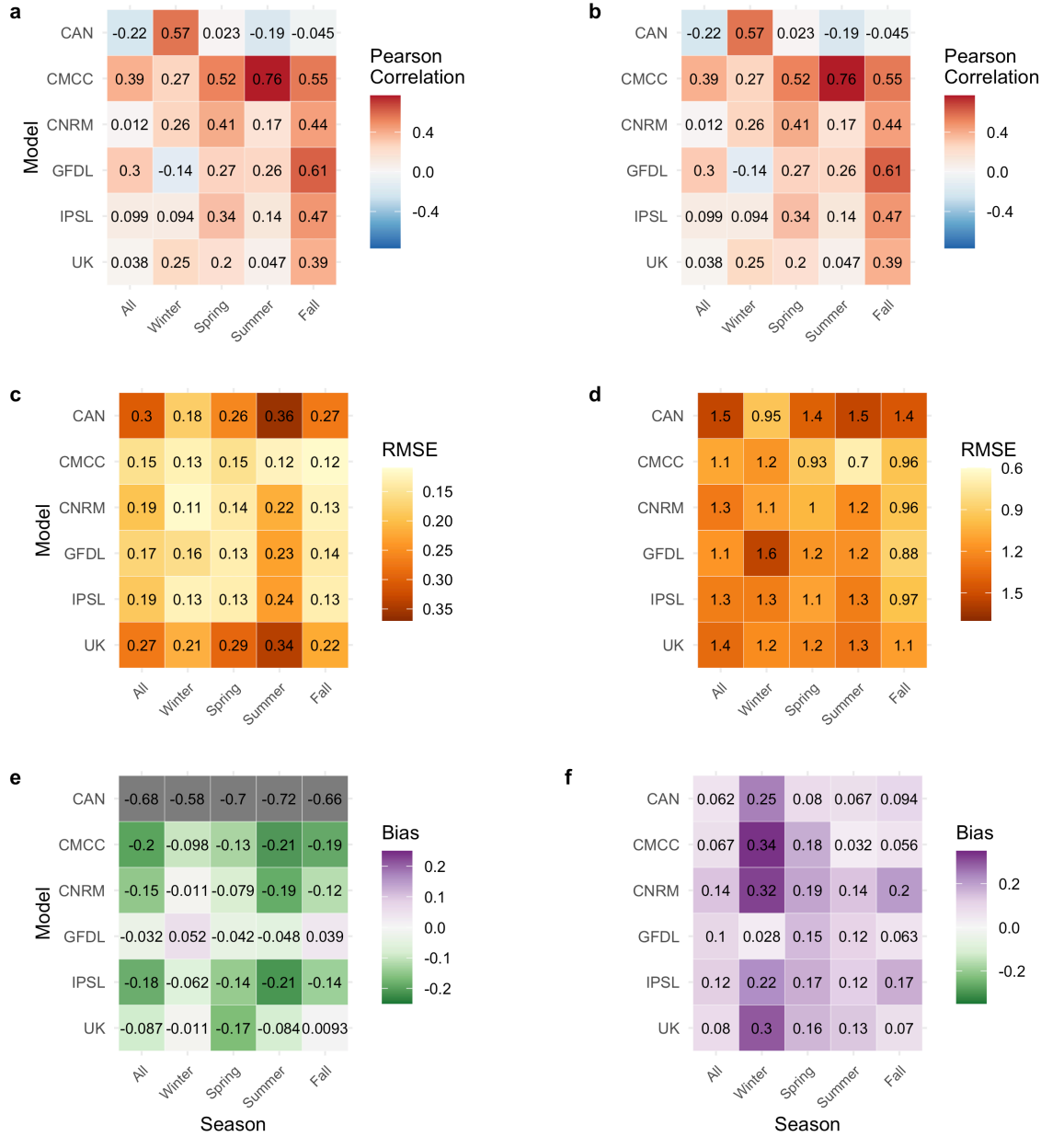
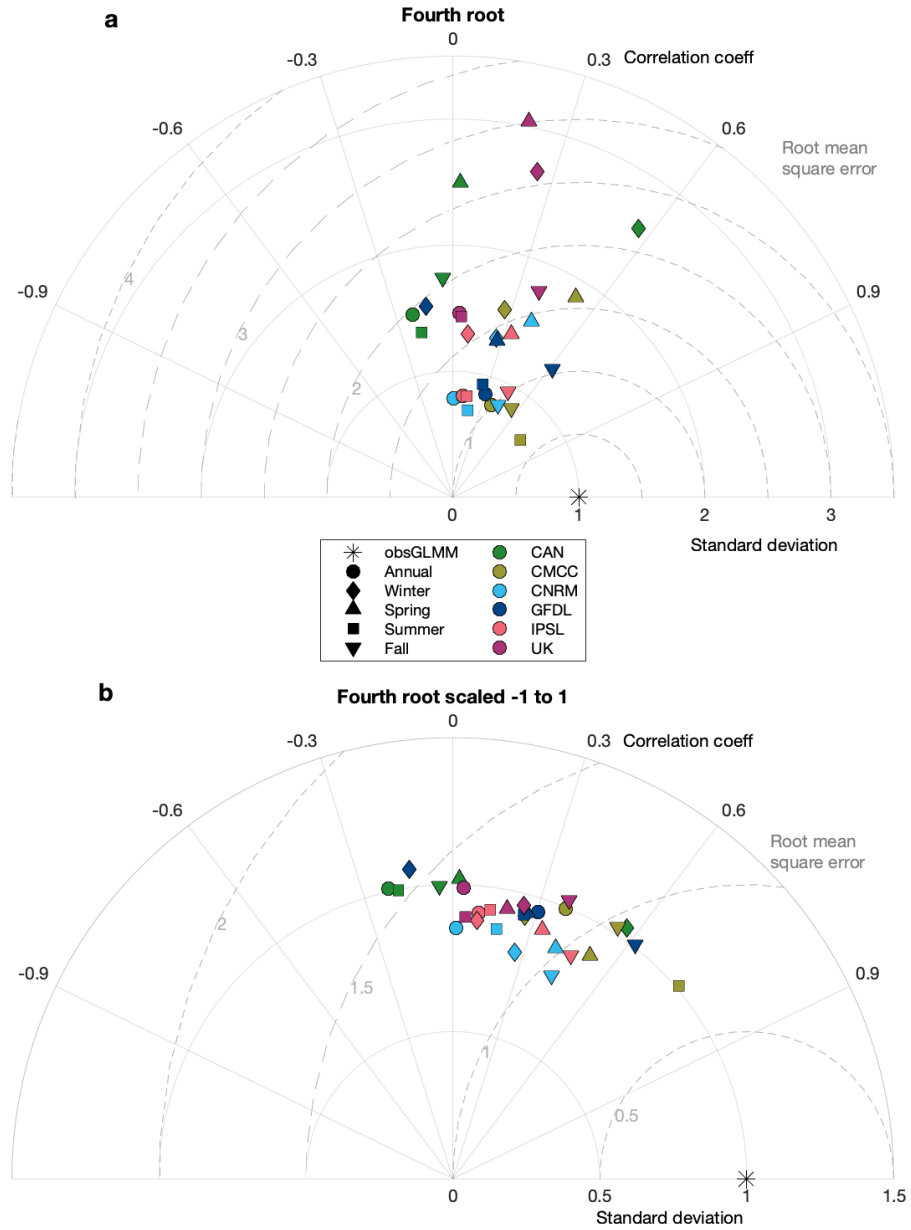


Figure 2. Pearson correlation coefficient, unbiased root mean square error (RMSE), and bias of mesozooplankton in the Historical simulations compared to the obsGLMM (left) without scaling and (right) scaled from -1 to 1. All biomasses were 4th-root transformation before further scaling. Southern hemisphere seasonal climatologies were shifted 6 months (e.g. Win = N Hem DJF and S Hem JJA). Grey shading denotes values outside the range of the colorbar.

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 3a). CAN and UK are often outliers on the Taylor diagram due to their large errors (Figure 2c) and standard deviations (Figure 3a). 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).

Using a scaling from -1 to 1 reveals consistent patterns by season and weaker patterns across the models (Figure 2b,d,f). The fall tends to have the lowest errors, though errors are similar across seasons and models (Figure 2d). The winter has the highest biases, which is in contrast to the non-scaled bias where winter climatologies produced the lowest (Figure 2f). In the Taylor diagram with scaling it is clear that the fall climatologies perform better relative to other seasons (Figure 2b,d, Figure 3b). Though the scaling resulted in normalized standard deviations near 1, CAN still appears as an outlier and UK also has lower skill than the other models due to their higher root mean square errors (Figure 3b).

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 4a). CAN chlorophyll *a* is oddly higher than the other ESMs and the obsGLMM in the LC biome (Figure 4b). CAN SST is cooler than the others in the LC biome, but higher in both HC biomes (Figure 4c). 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 vs. resources on the mesozooplankton biomass since chlorophyll *a* varies more across the ESMs (Figure 4b). Seasonal cycles of chlorophyll *a* vary both in peak magnitude and timing (Figure 4b). 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 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 4a). In individual biomes, the range of mesozooplankton biomass varies



(e.g. Win = N Hem DJF and S Hem JJA).

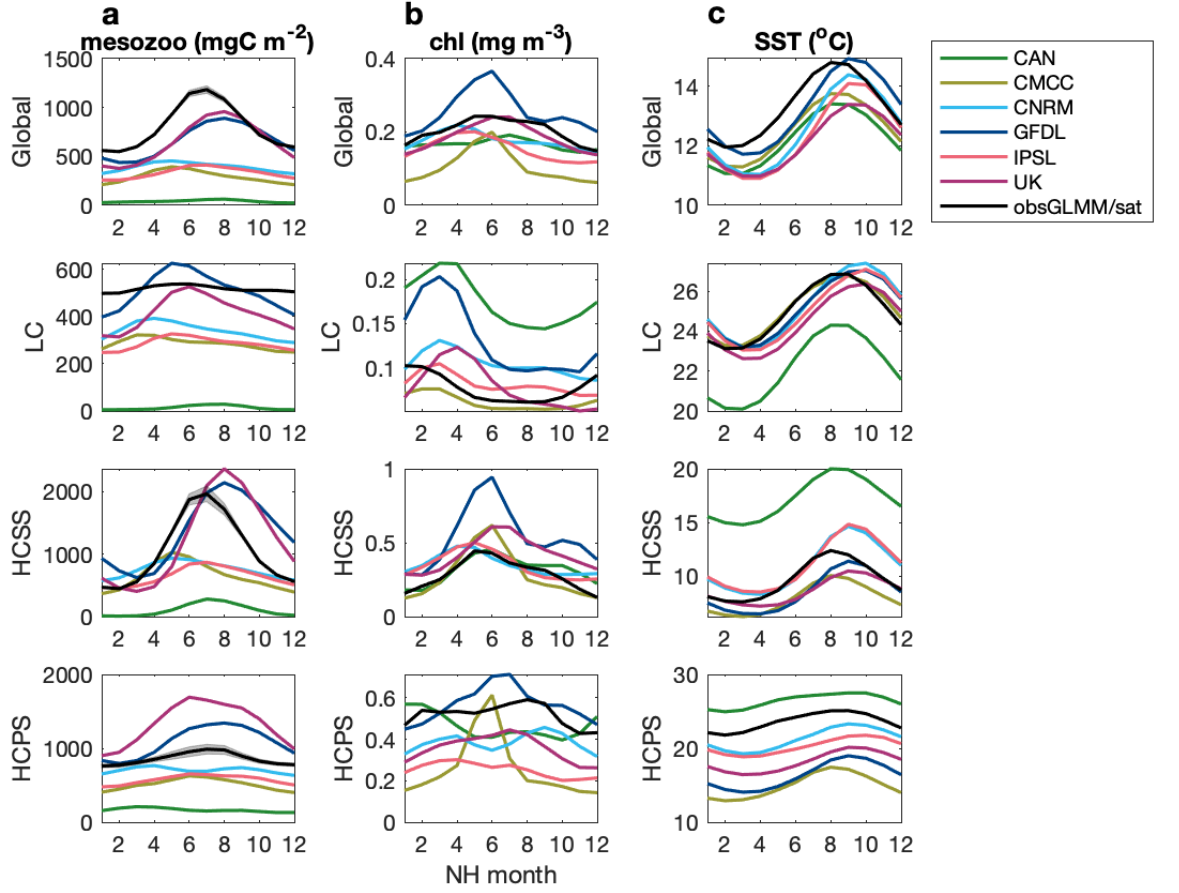


Figure 4. Annual climatologies of Historic (a) mesozooplankton biomass (mg C m^{-2}) in ESMs and obsGLMM. (b) surface chlorophyll *a* concentration (mg m^{-3}), and (c) sea surface temperature (SST) in ESMs and satellites (input to obsGLMM). Southern hemisphere seasonal climatologies were shifted 6 months (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 obsGLMM is ± 2 standard deviations.

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* 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 5). 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 5a). The percent change in mesozooplankton biomass over time roughly mimics that of surface chlorophyll *a* (Figure 5b) for the CAN and IPSL models, but the relationship between mesozooplankton with chlorophyll *a* or temperature (Figure 5c) 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 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 (Figure 5d-f).

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 (Figure 6a-b). All ESMs had steeper relationships than the obsGLMM globally, as well as by biome (Table S3). 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 COPEPOD-based constraints, the obsCS and obsCM, fell in the center between the obsGLMM and obsSM, with the three well constrained models clustering tightly around them.

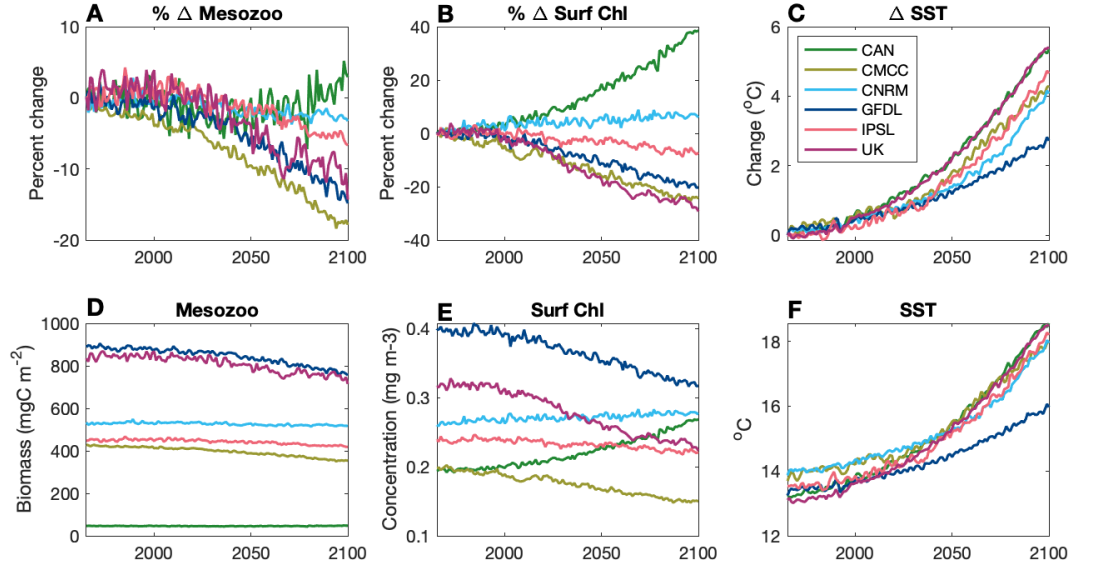


Figure 5. Globally averaged time series of ESM simulated (a,d) mesozooplankton biomass, (b,e) surface chlorophyll concentration, and (c,f) SST from 1965 to 2100. Shown are (a-c) percent change (mesozoo, chl) and change (SST) from 1951 and (d-f) absolute values.

From the biome perspective, the tightest relationship between ESM mesozooplankton biomass and chlorophyll *a* is found in the LC biome (Figure S4). In this biome, the obsGLMM and obsSM have slopes of 0.41 and 0.84, respectively, with the GFDL model having the only slope (0.56; Table S3, Figure 6a-b) that fell within the bounds of the observational products. The other models exhibited slopes between 0.95 and 2.02 (Table S3). 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 obsSM (Table S3). However, the scatter around these linear relationships (Figure S4) denotes the greater variability at the regional level beyond biomes, which differs by ESM (Figure 7a).

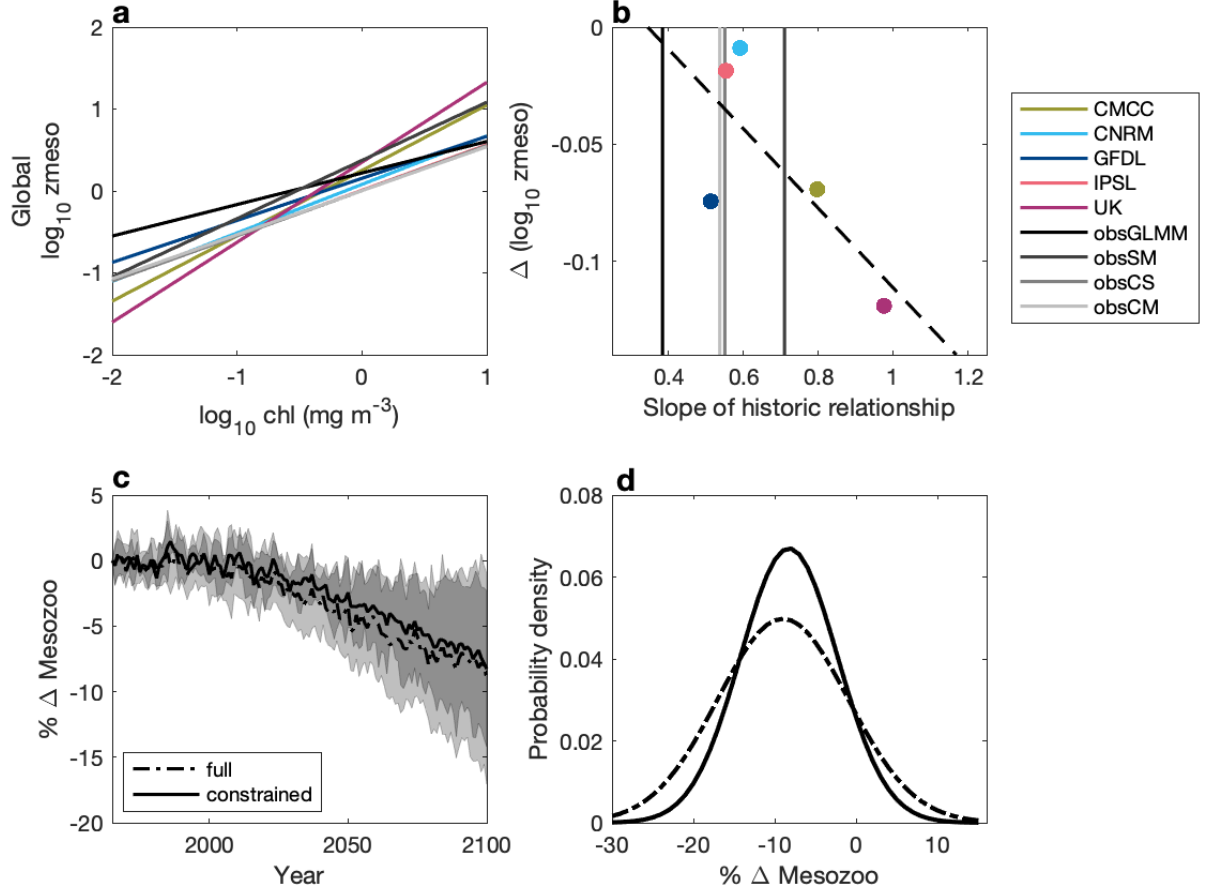


Figure 6. (a) Linear regressions of mesozooplankton biomass ($\log_{10} \text{ mg C m}^{-2}$) and surface chlorophyll a concentration ($\log_{10} \text{ mg m}^{-3}$) globally during the Historic period (1965-2014 mean). (b) Change in projected mesozooplankton biomass under SSP5.8-5 (2051-2100 mean) as a function of the relationship between mesozooplankton and chlorophyll under Historic 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 6 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).

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 strongly related to the historical relationship between mesozooplankton and surface chlorophyll a , where models that have a steeper mesozooplankton-chlorophyll a scaling slope exhibit greater declines in mesozooplankton biomass under climate change (Figure 6b). The global climate sensitivity of mesozooplankton biomass is primarily driven by the LC biome, which has greater changes in mesozooplankton biomass and stronger mesozooplankton-chlorophyll a relationships than the HC biomes (Table S3, Figure 7c). This mesozooplankton-chlorophyll a scaling thus provides an emergent constraint by which future projections of mesozooplankton biomass change can be slightly narrowed (Figure 6c-d).

4. DISCUSSION

Zooplankton have not been considered previously in major comparisons of the skill of ESMs (Bopp et al., 2013; Kwiatkowski et al., 2020; Séférian et al., 2020; Fu et al., *in revision*). 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 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 6a), 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 6b). 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 (Figure 6c-d).

4.1 Global biomass distributions in ESMs

Most ESMs produced reasonable estimates of mesozooplankton biomass. Mesozooplankton biomasses from all ESMs, except for the CAN model, were positively correlated with the observation-based statistical model (obsGLMM), although correlations were

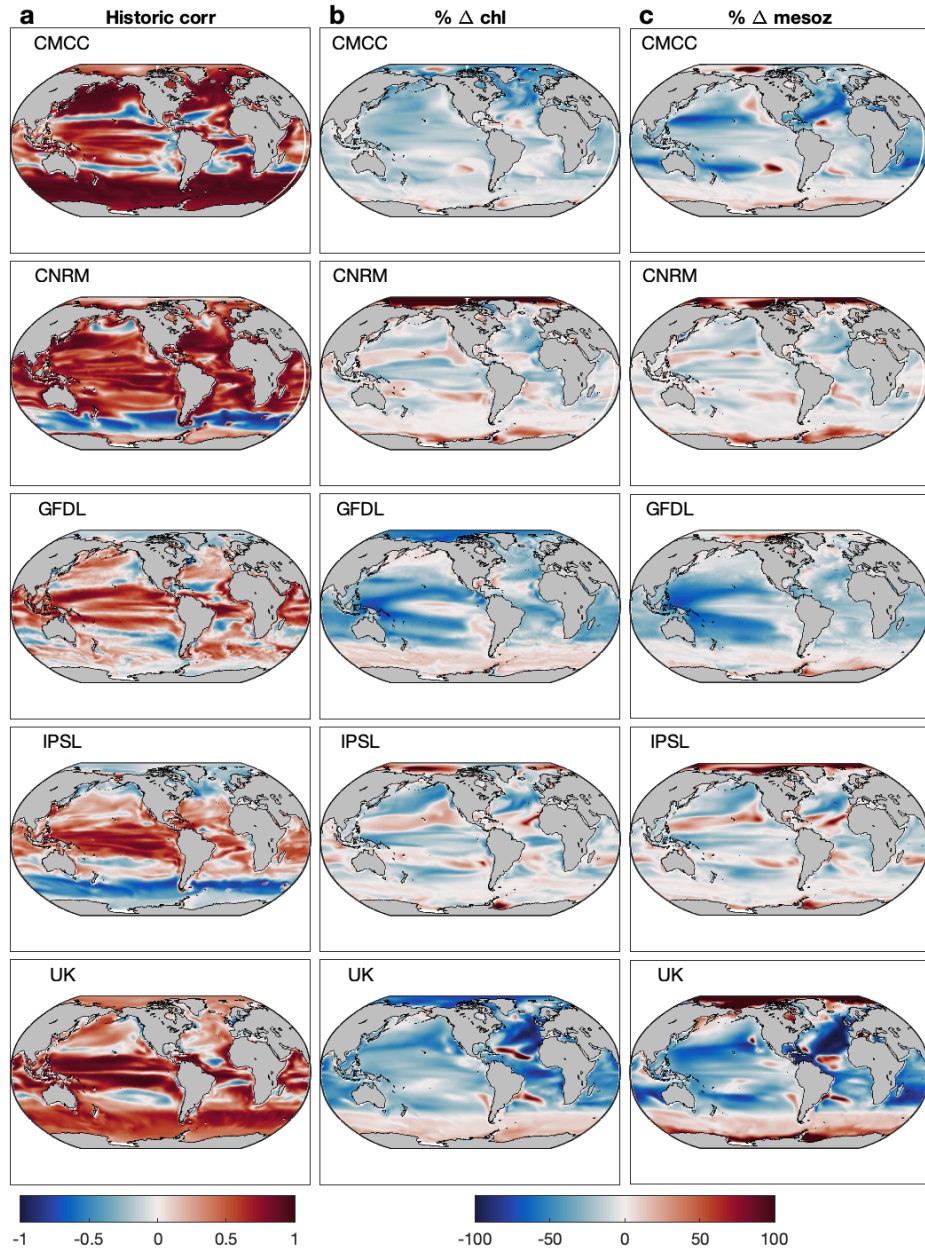


Figure 7. (a) Correlations between the 50-yr mean of mesozooplankton biomass and surface chlorophyll a concentration during the Historic period for each ESM. Percent change in mean (b) surface chlorophyll a concentration, and (c) mesozoo-

plankton biomass in the ESMs in the SSP5-8.5 (2051-2100) simulation compared to the Historical (1965-2014) simulation.

relatively weak (<0.5) with a few exceptions during particular seasons. Model error (root mean square error and bias) was greatest in summer and lowest in winter, whereas the highest correlations (greatest 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 provide insight into key elements of Earth system studies, namely Earth system predictability and prediction experiments. Because of the high model skill in the fall, experiments targeting mesozooplankton (among other quantities) should be initialized during the fall (e.g., Yeager et al., 2018; Park et al., 2019), as that season is one in which the models most closely match the observed ocean. This may have implications for evaluating fisheries predictions (Stock et al., 2017; Park et al. 2019).

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, 2012) span 2.38 orders of magnitude.

Reproduction of the contrast in oligotrophic vs. 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.

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 mesozooplankton and surface chlorophyll *a* (Richardson & Schoeman, 2004) as chlorophyll *a* can serve as a proxy for both phytoplankton biomass and primary production (Friedland et al., 2012; Marañón et al., 2014; Brewin et al., 2015). This was generally true, particularly

on the large scale, though there were regional exceptions. Both CMCC and UK had negative relationships with chlorophyll *a* in the HCPS biome under historic conditions (Figure 6). There were also negative relationships in polar regions (Figure 7). CNRM and IPSL had strong negative correlations with chlorophyll *a* in the Southern Ocean, while IPSL and GFDL had weaker negative correlations in the Arctic.

There are several potential explanations for the variable and sometimes negative relationships between mesozooplankton biomass and chlorophyll *a*. First, negative relationships could signify top-down control of phytoplankton biomass (chlorophyll *a* concentration) by mesozooplankton (Behrenfeld & Boss, 2014; Friedland et al., 2016). Second, it could reflect the disconnect between the longer turnover timescales of mesozooplankton (especially at cold temperatures) and the seasonal productivity window of phytoplankton. Since mesozooplankton biomass is lowest in winter, mesozooplankton in most regions are not sufficiently abundant to graze down the spring/summer phytoplankton bloom (e.g. Fennel & Neumann, 2015). This is exacerbated in ESMs that do not simulate diapause and other overwintering behaviors of zooplankton that cause stronger grazing pressure during phytoplankton blooms in the real world (Evans & Parslow, 1985). Third, ice dynamics of each ESM will also affect polar regions, since ice affects the light-harvesting ability of phytoplankton and may select for phytoplankton types with higher chlorophyll:carbon ratios (Sathyendranath et al., 2009). Thus, phytoplankton composition and thus the types of prey available to mesozooplankton will impact the mesozooplankton-chlorophyll relationship in all regions for each ESM. Fourth, relationships between mesozooplankton and chlorophyll *a* will be affected by different parameterizations of the influence of temperature on physiological rates in each model. Last, the negative relationships between mesozooplankton and chlorophyll *a* could be influenced by the inclusion of arctic regions in the HCPS biome. The difference in light seasonality in HCPS regions over wide latitudes could confuse the signal by combining multiple mesozooplankton-chlorophyll *a* relationships into one (Figure S4).

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 6b). 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 8a) or globally (CNRM, Figure 5b). These are regions where the size structure of phytoplankton should shift towards 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 6a, Table S3), and chlorophyll *a* declines much more in projections (Figures 5b, 7b). 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). 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 5a, 6b), which is further illustrated by the greater percent changes in projected mesozooplankton biomass compared to chlorophyll (Figure 7b-c). 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 4) and long-term projections (Figure 5). 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 to 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 1.88-2.0, 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 5b). 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 5a). 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 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 a key driver for 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 (Kwiatkowski et al., 2017; Eyring et al., 2019; Hall et al., 2019). Such analyses isolate a relationship within individual ESMs

that demonstrates the same response to climate change across ESMs. Ensemble members could be removed from an ensemble (or downweighted) 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. We were able to identify a positive relationship between historic mesozooplankton biomass and surface chlorophyll *a* concentration that spanned trophic gradients in observations and two observation-based products. 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. Thus, ensembles of mesozooplankton biomass could be reduced to CNRM, GFDL, and IPSL model output (Figure 6c-d). 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 (Figure 6c-d), as these three models still encompassed a large range of the climate response of mesozooplankton (Figure 5a). 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. This global emergent constraint is strongly influenced by the oligotrophic, LC regions where chlorophyll *a* generally decreases under climate change (Figure 7b), resulting in larger reductions of mesozooplankton in the models where mesozooplankton are more closely tied to chlorophyll *a* (Figure 7c). This pattern is accordant with the processes of warming and stratification causing the greatest declines in large phytoplankton biomass and thus mesozooplankton biomass in the oligotrophic gyres. The link between the mesozooplankton-chlorophyll relationship and the change in mesozooplankton biomass under climate change could serve as a significant emergent constraint on climate change projections of plankton dynamics.

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 (Strömberg et al., 2009; Buitenhuis et al., 2010; Moriarty & O’Brien, 2012; Heneghan et al., 2020), more assessment of zooplankton becomes possible. In addition to zooplankton biomass, model developers could also consider assessment of other terms, such as physiological rates (Hirst et al., 2003; Kiørboe & Hirst, 2013; Heneghan et al., 2020), their associated model parameters (Buitenhuis et al., 2006; Stock et al., 2014b; Rohr et al., *submitted*), 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. Hjøllø 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 $<1^\circ$ 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 (Stock et al., 2011; Drenkard et al., 2021). 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 (Stock et al., 2011, Drenkard et al. 2021) 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; Heneghan et al., *submitted*).

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. 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

The obsGLMM is available at https://github.com/MathMarEcol/CMIP6_ZoopStatisticalModel. All other observations and model outputs are available at the websites cited in the text of their descriptions.

ACKNOWLEDGEMENTS

We thank CA Stock and T Rohr for helpful comments on earlier drafts, and O Aumont, M Butenschön, J Christian, T Lovato, R Séférian, and A Yool for clarifications on their BGC sub-models. 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: Cambridge University Press.
- Arrhenius, S. (1889). On the reaction rate of the inversion of non-refined sugar upon souring. *Z. Phys. Chem.*, 4, 226–248.
- Arhonditsis, G. B., & Brett, M. T. (2004). Evaluation of the current state of mechanistic aquatic biogeochemical modeling. *Mar. Ecol. Prog. Ser.*, 271,

13–26. <https://doi.org/10.3354/meps271013>

Aumont, O., Ethé, C., Tagliabue, A., Bopp, L., & Gehlen, M. (2015). PISCES-v2: An ocean

biogeochemical model for carbon and ecosystem studies. *Geosci. Model Dev.*, 8, 2465–2513.

<https://doi.org/10.5194/gmd-8-2465-2015>

Banase, 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). New York, NY: Plenum Press.

Behrenfeld, M. J. & Boss, E. S. (2014). Resurrecting the ecological underpinnings of ocean plankton blooms. *Annual review of marine science*, 6, 167–194.

Behrenfeld, M. J. & Falkowski, P. G. (1997). A consumer’s guide to phytoplankton primary productivity models. *Limnology and oceanography*, 42(7), 1479–1491.

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

Bode, A., Álvarez-Ossorio, M. T., & González, N. (1998). Estimations of mesozooplankton biomass in a coastal upwelling area off NW Spain. *J. Plankton Res.*, 20, 1005–1014. <https://doi.org/10.1093/plankt/20.5.1005>

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, 127–135.

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.

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.

Buitenhuis, E. T., Rivkin, R. B., Sailley, S. & Le Quéré, C. (2010). Biogeochemical fluxes through microzooplankton. *Global biogeochemical cycles*, 24(4), GB4015.

- Calbet, A. (2001). Mesozooplankton grazing effect on primary production: a global comparative analysis in marine ecosystems. *Limnology and Oceanography*, 46(7), 1824-1830.
- Calbet, A. & Landry, M. R. (2004). Phytoplankton growth, microzooplankton grazing, and carbon cycling in marine systems. *Limnology and Oceanography*, 49(1), 51-57.
- 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 Res. Part II Top. Stud. Oceanogr.*, 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*, pp.1-68.
- Conley K. R., Lombard F., & Sutherland K. R. (2018). Mammoth grazers on the ocean's minuteness: a review of selective feeding using mucous meshes. *Proc. R. Soc. B*, 285, 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*. 1969-1987.
- Eppley, R. W. (1972). Temperature and phytoplankton growth in the sea. *Fish. Bull.*, 70, 41063-41085.
- Evans, G. T., & Parslow, J. S. (1985). A Model of Annual Plankton Cycles. *Biological Oceanography*, 3(3), 327-347.
- 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.
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., et al. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.*, 9, 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. *Nat. Clim. Chang.*, 9, 102-110. <https://doi.org/10.1038/s41558-018-0355-y>

- Fennel, W., & Neumann, T. (2014). *Introduction to the Modelling of Marine Ecosystems*. Amsterdam: Elsevier Science.
- 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. *J Plankton Res*, *32*, 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), p.e28945.
- Friedland, K. D., Record, N. R., Asch, R. G., Kristiansen, T., Saba, V. S., Drinkwater, K. F., et al. (2016). Seasonal phytoplankton blooms in the North Atlantic linked to the overwintering strategies of copepods. *Elementa: Science of the Anthropocene*, *4*, 000099.
- 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*, 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.
- 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>
- Fu, W., Moore, J. K., Primeau, F., Collier, N., Ogunro, O. O., Hoffman, F. M., & Randerson, J. T. (*in revision*). Evaluation of CMIP models with IOMB: Rates of contemporary ocean carbon uptake linked with vertical temperature gradients and transport to the ocean interior. *JGR Oceans*.
- 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.
- 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*, 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. *Nat. Clim. Chang.*, *9*, 269–278.

<https://doi.org/10.1038/s41558-019-0436-6>

Hansen, P. J., Bjørnsen, P. K. & Hansen, B. W. (1997). Zooplankton grazing and growth: Scaling within the 2-2,- m body size range. *Limnology and oceanography*, 42(4), 687-704.

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*. <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. F., Everett, J. D., Sykes, P., Batten, S. D., Edwards, M., Takahashi, et al. (2020). A functional size-spectrum model of the global marine ecosystem that resolves zooplankton composition. *Ecological Modelling*, 435, 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.

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

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>.

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.

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.

Ikeda, T. (1985). Metabolic rates of epipelagic marine zooplankton as a function of body mass and temperature. *Marine Biology*, 85(1), 1-11.

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.

Kjørboe, 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.

- 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.
- Kwiatkowski, L., Aumont, O., & Bopp, L. (2019). Consistent trophic amplification of marine biomass declines under climate change. *Glob. Change Biol.*, 25, 218–229. <https://doi.org/10.1111/gcb.14468>
- Kwiatkowski, L., Torres, O., Bopp, L., Aumont, O., Chamberlain, M., Christian, J., et al. (2020). Twenty-first century ocean warming, acidification, de-oxygenation, and upper ocean nutrient decline from CMIP6 model projections. *Biogeosciences Discuss.*, 1–43. <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, 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.
- 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, e2021GL094367. <https://doi.org/10.1029/2021GL094367>
- Longhurst, A. (1995). Seasonal cycles of pelagic production and consumption. *Progress in Oceanography*, 36, 77–167.
- López-Urrutia, Á., San Martín, 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.
- 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.
- 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*, e2021MS002814.
- Marañón, E., Cermeño, P., Huete-Ortega, M., López-Sandoval, D. C., Mourinho-Carballido, B. & Rodríguez-Ramos, T. (2014). Resource supply overrides temperature as a controlling factor of marine phytoplankton growth. *PloS one*, 9(6), e99312.
- Maier-Reimer, E. & Hasselmann, K. (1987). Transport and storage of CO₂ in

- the ocean—an inorganic ocean-circulation carbon cycle model. *Climate Dynamics*, 2(2), 63-90.
- 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, 113–117
- Moriarty, R. & O’Brien, T. D. (2012). Distribution of mesozooplankton biomass in the global ocean. *Earth System Science Data*, 5(1), 45-55.
- 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.
- 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.
- Postel, L. (1990). The mesozooplankton response to coastal upwelling off West Africa with particular regard to biomass. *Mar. Sci. Rep.*, 1, 1–127.
- Postel, L., Fock, H. & Hagen, W. (2000). Biomass and abundance. In R. Harris, P. Wiebe, J. Lenz, H. R. Skjoldal, M. Huntley, *ICES Zooplankton Methodology Manual* (pp. 83-192). Cambridge, MA: Academic Press.
- Richardson, A. J., & Schoeman, D. S. (2004). Climate Impact on Plankton Ecosystems in the Northeast Atlantic. *Science*, 305, 1609–1612. <https://doi.org/10.1126/science.1100958>
- Rohr, T., Richardson, A. J., Lenton, A., Shadwick, E., & Chamberlain, M. (*submitted*) Recommendations for the Formulation of Grazing in Marine Biogeochemical and Ecosystem Models. *Progress in Oceanography*.
- Ryther, J. H. (1969). Photosynthesis and fish production in the sea. *Science*, 166(3901), 72-76.
- 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.
- Sameoto, D., P. Wiebe, J. Runge, L. Postel, J. Dunn, C. Miller, & S. Coombs. (2000). Collecting zooplankton. In R. Harris, P. Wiebe, J. Lenz, H. R. Skjoldal, M. Huntley, *ICES Zooplankton Methodology Manual* (pp. 55-81) Cambridge, MA: Academic Press.
- Sarmiento, J. L., Orr, J. C. & Siegenthaler, U. (1992). A perturbation simulation of CO₂ uptake in an ocean general circulation model. *Journal of Geophysical Research: Oceans*, 97(C3), 3621-3645.
- 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*

- Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, 106(D7), 7183-7192.
- Terhaar, J., Frölicher, T. L. & Joos, F. (2021). Southern Ocean anthropogenic carbon sink constrained by sea surface salinity. *Science Advances*, 7(18), p.eabd5964.
- 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.
- 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*, 1-9.
- Wiebe, P. H. (1988). Functional regression equations for zooplankton displacement volume, wet weight, dry weight, and carbon : a correction. *Fish. Bull.*, 86, 833-835.
- Wiebe, P. H. & Benfield, M. C. (2003). From the Hensen net toward four-dimensional biological oceanography. *Progress in Oceanography*, 56(1), 7-136.
- 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.
- 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.
- 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. *Geosci. Model Dev.*, 6, 1767-1811. <https://doi.org/10.5194/gmd-6-1767-2013>
- 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.