Pathways for avoiding ocean biogeochemical damage: Mitigation limits, mitigation options, and projections

Timothée Bourgeois[1], Olivier Torres[2], Friederike Fröb[3], Aurich Jeltsch-Thömmes[4], Giang T. Tran[5], Jörg Schwinger[1], Thomas L. Frölicher[4,6], Jean Negrel[1], David Keller[5], Andreas Oschlies[5], Laurent Bopp[2], Fortunat Joos[4]

[1]NORCE Climate & Environment, Bjerknes Centre for Climate Research, Bergen, 5005, Norway
[2]LMD-IPSL, CNRS, Ecole Normale Supérieure/PSL Res. Univ, Ecole Polytechnique, Sorbonne Université, Paris, 75005, France
[3]Geophysical Institute, University of Bergen, Bjerknes Center for Climate Research, Bergen, 5005, Norway
[4]Climate and Environmental Physics, Physics Institute, University of Bern, Bern, 3012, Switzerland
[5]Marine Biogeochemical Modelling, GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel, Kiel, 24105, Germany
[6]Oeschger Centre for Climate Change Research, University of Bern, Bern, 3012, Switzerland

Correspondence to: Timothée Bourgeois (tbou@norceresearch.no)

Abstract. Anthropogenic greenhouse gas emissions cause multiple changes in the ocean and its ecosystems through climate change and ocean acidification. These changes can occur progressively with rising atmospheric carbon dioxide concentrations, but there is also the possibility of large-scale abrupt, and/or potentially irreversible changes, which would leave limited opportunity for marine ecosystems to adapt. Such changes, either progressive or abrupt, pose a threat to biodiversity, food security, and human societies. However, it remains notoriously difficult to determine exact limits of a "safe operating space" for humanity. Here, we map, for a variety of ocean impact metrics, the crossing of limits, which we define using the available literature and to represent a wide range of deviations from the unperturbed state. We assess the crossing of these limits in three future emission pathways: two climate mitigation scenarios, including an overshoot scenario, and one high-emission no-mitigation scenario. These scenarios are simulated by the latest generation of Earth system models and large perturbed-parameter ensembles with two Earth system models of intermediate complexity. Using this comprehensive model database, we estimate when and at which warming level 4 mitigation limits based on expert judgement for 14 different impact metrics are exceeded along with an assessment of uncertainties. We find that under the high-emissions scenario, the two highest limits are exceeded with high confidence for the marine heatwaves' duration, expansion of ocean areas that are undersaturated with respect to aragonite, decreases in plankton biomass, and loss of Arctic summer sea ice extent. The probability of exceeding a given limit generally decreases clearly under low-emissions scenario. Yet, exceedance of ambitious limits related to Arctic aragonite undersaturation, plankton biomass, and Arctic summer sea ice extent are projected to be difficult to avoid (high confidence) even under the low-emissions scenario. The scenario including a temporary overshoot reduces with high confidence the risk of exceeding mitigation limits by year 2100 related to the marine heatwave duration, metabolic index, plankton biomass, Atlantic meridional overturning circulation, aragonite undersaturation, global deoxygenation, and Arctic

summer sea ice extent compared to the high-emissions scenario. Our study highlights the urgent need for ambitious mitigation efforts to minimize extensive impacts and potentially irreversible changes to the world's oceans and ecosystems.

1 Introduction

Earth system models (ESMs) are invaluable tools to simulate the climate outcomes of future emission pathways. However, due to computational constraints, ESMs are limited in terms of spatial resolution and complexity of represented processes (Chen et al., 2021). Consequently, it remains notoriously difficult to assess climate change impacts that occur at smaller scales or in systems that are not exhaustively represented in ESMs, for example, marine ecosystems. Large-scale changes in important drivers of marine ecosystem processes (for example, warming, deoxygenation, and acidification) are often taken as a measure of potential ecosystem damage ("ecosystem stressors"). As such changes will usually occur simultaneously, the term "multiple potential ecosystem stressors" has been coined to describe the threat that climate change poses to marine ecosystems (Bopp et al., 2013; Gattuso et al., 2015; Gruber, 2011; Kwiatkowski et al., 2020).

Although it is generally well-known which climate variables will have an adverse impact on ecosystems and societies if they are altered by human activity, exact limits that should not be exceeded are often difficult to define. This might be the case either because impacts occur gradually with changes in a driver variable (and it thus remains an ethical or economic question how much damage can be accepted), or because a limit exists but is uncertain. The latter will be the case if tipping points exist in the system, a crossing of which will lead to large and irreversible changes (e.g., Lenton et al., 2008; Armstrong McKay et al., 2022). For ecosystems in particular, the possibility exists that gradual changes in the physical or biogeochemical state lead to the crossing of tipping points (Heinze et al., 2021). Nevertheless, our knowledge on the impacts of these changes on marine ecosystems is growing. Thermal changes induced by global warming are altering the productivity of phytoplankton functional types and reshape established interspecific competition in marine ecosystems (Kordas et al., 2011; Dutkiewicz et al., 2013; Anderson et al., 2021). The shoaling of the calcium carbonate ($CaCO_3$) saturation horizon due to ocean acidification is threatening calcifying organisms (Orr et al., 2005; Doney et al., 2020). Ocean deoxygenation and the expansion of oxygen minimum zones contribute to marine aerobic habitat loss (Diaz and Rosenberg, 2008; Pinsky et al., 2020; Morée et al., 2023; Fröb et al., 2024). Shifting circulation patterns affect fish migrations and human societies (Van Gennip et al., 2017; Schwinger et al., 2022). Finally, upper-ocean stratification changes alter ocean primary productivity and community structures by exacerbating surface nutrient depletion (e.g., Fu et al., 2016).

In this study, we follow the approach of Steinacher et al. (2013) and define a set of 14 impact metrics associated to 4 mitigation limits. We aim at determining the probability of staying within a given mitigation limit based on scenario simulations from state-of-the-art Earth system models from the latest Coupled Model Intercomparison Project (CMIP6), while earlier studies addressing multiple limits relied on perturbed parameter ensembles from Earth system models of intermediate complexity (EMICs) (e.g., Steinacher et al., 2013; Battaglia and Joos, 2018). As an original aspect of our study, we also include results from perturbed parameter ensembles of two EMICs for comparison with our CMIP6 model ensemble.

2 Method

2.1 Definition of impact metrics and mitigation limits

We consider 14 illustrative impact metrics, and each of them are associated with 4 mitigation limits (Table 1). Mitigation limits are ordered according to the expected severity of impacts when exceeding the limit, that is, exceeding mitigation limit 4 for a given metric is expected to result in more severe impacts than exceeding mitigation limit 1, and correspondingly, staying under limit 1 is more ambitious because a higher emission reduction would be needed to stay below this limit. Mitigation limits at a given level are not necessarily dependent, i.e., they can be exceeded at different time and global warming levels.

A literature review combining observations with simulation studies on critical limits and thresholds in the ocean system is conducted to define the impact metrics and mitigation limits. While the aim of this study is to define mitigation limits based as much as possible on the literature, many metrics suffer from a lack of knowledge regarding the assessment of actual impacts that an exceedance would have on the Earth system or ecosystem functioning. Some physical metrics have been more thoroughly investigated, while biogeochemical metrics are less constrained. Observations and laboratory experiments suggest numerous critical limits for key ecosystem stressors. Moreover, these limits are species-dependent and can vary over a wide range. Thus, for some metrics, we favour illustrative limits for relative changes to characterise reasonably safe levels instead of absolute changes. These choices could be refined with future research and further dialogue with stakeholders. If the literature does not permit to define mitigation limits for a specific metric, then we use ad hoc limits that cover the simulated mitigation space from very strong mitigation to very little or no mitigation effort.

The impact metrics include five physical parameters, related to surface atmospheric warming, marine heatwaves, steric sea level rise, sea ice extent, and the Atlantic Meridional Overturning Circulation (AMOC), five chemical parameters, related to global and regional ocean acidification and deoxygenation, and four ecosystem parameters, related to productivity, biomass, organic matter export, and metabolic performance.

Physical metrics

Global mean surface air temperature (SAT) is an important metric of the climate system and has strong and direct influences on ecosystems as well as human systems, i.e., many other important indicators and metrics co-vary with temperature. We pick the mitigation limits of 1.5°C and 2°C increase since the 1850-1900 mean based on the Paris agreement (UNFCCC, 2015). Two additional limits of 3°C and 4°C represent temperature limits beyond which severe impacts and the triggering of global tipping elements could be possible (Armstrong McKay et al., 2022; Masson-Delmotte et al., 2021). These limits for global mean SAT increase have also been previously used in Steinacher et al. (2013). We consider marine heatwaves due to their strong global and regional impact on marine ecosystems (Oliver et al., 2021). We define a marine heatwave day as the local daily mean sea surface temperature (CMIP6 variable *tos*) above the 90th percentile relative to a fixed seasonally varying 1850-1900 baseline. Then we average globally the duration over the year. Because such global metric is poorly constrained by

observations, we spread evenly the mitigation limit values over the year as 90, 180, 270, and 360 (i.e., permanent heatwave) days.

The third physical metric chosen is the rise of steric sea level (SSL). Sea level is rising and accelerating, which poses a significant challenge to coastal ecosystems and community livelihoods. We are only considering the SSL rise because the models used in our study do not simulate the melting of glaciers and ice sheets. While strongly connected to global warming, SSL rise shows a delayed response due to the various time scales on which its components operate. The SSL rise is approximated to be 40 % of the total sea-level rise of 0.2 m today (Church et al., 2011; Fox-Kemper et al., 2021; WCRP Global Sea Level Budget Group, 2018). O'Neill et al. (2017) estimate that risks related to SSL rise are at a moderate level at about 0.1 m above the 1986-2005 level, and transition to high risks are expected at around 1 m above the same reference level. Hinkel et al. (2014) find that under no adaption, 0.25-1.23 m of global sea-level rise (i.e., 0.1 to 0.5 m of SSL rise assuming a constant steric to sea level rise ratio) in 2100 would expose 0.2 - 4.6 % of the global population to flooding annually. Hermans et al. (2021) found a mean SSL rise of 0.27 m in 2100 simulated by CMIP5 and CMIP6 ensembles under high-emissions scenarios. Thus, we chose to define the four limits as 0.2, 0.3, 0.4, and 0.5 m relative to the period 1850-1900, to encompass the range found in the literature.

Changes in summer Arctic sea-ice extent have a direct impact on the climate system through the albedo feedback. Furthermore, a substantial reduction of Arctic sea ice could threaten the livelihood of organisms that depend upon habitats provided by sea ice. Arctic sea ice is projected to decline, and an ice-free summer state is a possibility even with a stabilised global warming of 1.5°C (~1 % chance of individual ice-free years by the end of the century; Pörtner et al., 2019). The ice-free state is defined as a September sea ice extent below 10^6 km². We further define three more limits up to 4 x 10^6 km² following the projected range from Stroeve et al. (2012) and Peng et al. (2020). The 4 x 10^6 km² limit has already been reached in 2012 and 2020.

A collapse of the Atlantic meridional overturning circulation (AMOC) is often considered a more distant tipping point (Lenton, 2012) even though recent literature estimated that we cannot rule out that AMOC is on course to collapse (Van Westen et al., 2024). The estimated probabilities from expert elicitation for the shutdown of AMOC (until 2100) is 0-0.2 for low (<2°C), 0-0.6 for medium (2-4°C), and 0.05-0.95 for high climate change (4-8°C), according to Zickfeld et al. (2007) and Kriegler et al. (2009). A weakening of the AMOC this century is expected (Pörtner et al., 2019; Fox-Kemper et al., 2021), which can cause, for example, changes in the rainfall, storm frequency in Northern Europe and a decrease in marine productivity in the North Atlantic (Pörtner et al., 2019). The four limits are chosen to cover a typical range of model responses (Weaver et al., 2012; Weijer et al., 2020). We compute the strength of the AMOC as the vertical maximum of the stream function at 26°N below 500 m depth.

Ocean acidification metrics

Ocean Ω aragonite (Ω_{arag}), or the level of saturation of the least-stable form of calcium carbonate in seawater, is a common indicator of the potential for biotic calcification. Ocean acidification could lead to undersaturation ($\Omega_{arag}<1$) and dissolution of calcium carbonate in parts of the surface ocean during the 21st century, which might have detrimental effects on marine

ecosystems (Orr et al., 2005). Studies have shown that no prominent present-day coral reefs exist in environments with $\Omega_{arag}<3$ (Guinotte et al., 2003; Hoegh-Guldberg et al., 2007; Kleypas et al., 1999). A lower limit of 1.5 has been used previously to signify water which may be stressful to larvae of shellfish such as oysters (Ekstrom et al., 2015; Gimenez et al., 2018). For $\Omega_{arag}<1.5$, marine organisms are believed to have trouble forming shells during the first few days of their life (Waldbusser et al., 2015). Guinotte et al. (2006) estimated that over 95 % of cold water bioherm-forming corals were found in areas that were supersaturated ($\Omega_{arag}>1$). We define three ocean acidification metrics in terms of area fractions. The first two metrics, A_{SO} and A_{Arctic}, are respectively the surface area fractions of the Southern Ocean (south of 50°S) and the Arctic Ocean (north of 70°N) undersaturated with respect to aragonite ($\Omega_{arag}<1$; annual mean), which means that seawater becomes corrosive to aragonitic shells of marine organisms (Doney et al., 2009; Fabry et al., 2009). The selected limits for these metric range from 20 % to 80 % following Steinacher et al. (2009, 2013). The third ocean acidification metric, $A_{\Omega>3}$, addresses areas with high saturation states ($\Omega_{arag}>3$) that are mainly found in the tropics and subtropics. We define this variable as the percentage of the global ocean surface area with $\Omega_{arag}>3$ that has been lost since pre-industrial times and select limits from 50 % to 100 % (Steinacher et al., 2013).

To account for carbonate chemistry biases in the present-day mean state simulated by the CMIP6 ESMs, we follow the methodology of Terhaar et al. (2020). Changes in aragonite saturation state (Ω_{arag}) have been computed offline using mocsy 2.0 (Orr and Epitalon, 2015) from regridded annual CMIP6 model outputs of dissolved inorganic carbon (DIC), alkalinity, sea water temperature (T), and sea water salinity (S). These modelled changes have then been added to the contemporary saturation state that we derived from the observation-based GLODAPv2 data product for DIC and alkalinity (Lauvset et al., 2016), and the World Ocean Atlas 2013 for T and S (Locarnini et al., 2013; Zweng et al., 2013).

Other biogeochemical metrics

Marine species have been observed to die after exposure to a wide range of critical O_2 levels, from 8.6 mg O_2 L^{-1} (ca. 275 µmol L^{-1}) to anoxia (Vaquer-Sunyer and Duarte, 2008). Critical O_2 levels are species- and stage-specific (Ekau et al., 2010), making it challenging to define common limits. Globally, dissolved O_2 is projected to decline by 1.81 to 3.45 % by year 2100 under CMIP5 representative concentration pathways (RCP) (Bopp et al., 2013). Subsurface (100-600 m) O_2 is projected to decline by 3.1-4 % under RCP8.5 and 0.1-0.5 % under RCP2.6. The projected decline in the subsurface (100-600 m) dissolved O_2 concentration for CMIP6 models under CMIP6 shared socioeconomic pathways (SSP) vary from -6.36 to -13.27 µmol L^{-1} by the end of the century (Kwiatkowski et al., 2020). The equivalent range for CMIP5 models under RCP scenarios is from -3.71 to -9.51 µmol L^{-1}. Due to large differences between models and when compared to observations, we decided to define relative limits for two metrics: mean global full-depth O_2 concentration and volume of hypoxic waters above 1000 m depth (i.e., waters with <63 µmol L^{-1}; Limburg et al., 2020).

The decline of marine net primary productivity (NPP) is considered one of the main stressors of open ocean ecosystems (Bopp et al., 2013). Mitigation limits for marine NPP are defined in terms of relative changes for the same reason as above. Kwiatkowski et al. (2020) shows changes from -0.56 ± 4.12 % under SSP1-2.6 to -2.99 % ± 9.11 % under SSP5-8.5 for CMIP6

models, while the range for CMIP5 models is from -3.42 ± 2.47 % under RCP2.6 to -8.54 ± 5.88 % under RCP8.5 until year 2100. Thus, we set the limits to 2, 3.5, 4 and 8 % relative to 1850-1900. In addition to NPP, we consider changes in plankton biomass (ΔBiomass) because projected plankton biomass has been considered as a more robust metric reflecting the impact of climate change on marine ecosystems (Bopp et al., 2022; Tittensor et al., 2021). The ΔBiomass metric represents the change in the sum of phytoplankton and zooplankton biomass (CMIP6 variables *zooc* and *phyc*) and its limits are the same as those for NPP.

We consider changes in the upper ocean metabolic index and changes in POM export between 30°S and 30°N as indicators of the compound effects of warming and oxygen changes on viable habitat and the survival of marine species (Battaglia and Joos, 2018). The metabolic index, Φ, is defined as the ratio of O_2 supply to an organism's resting O_2 demand. Warming ocean and lower partial pressure of O_2 is expected to reduce the globally averaged upper ocean metabolic index, which was shown to restrict viable habitats (Deutsch et al., 2015). Φ has been calculated using the mean ecophysiotype of the 61 species used in Fröb et al. (2024) and further described in Deutsch et al. (2020). The export of particulate organic material is the primary food source for deep-sea organisms. Thus, the POM export between 30°S and 30°N serves as an indicator of food availability in deep sea habitats. The limits of 4, 6, 8, and 10 % for these two indicators are based on the result of Battaglia and Joos (2018).

Table 1. Impact metrics and corresponding mitigation limits for changes until year 2100. Mitigation limits that are considered for ESMs only are marked with an asterisk.

Impact metric	Description	Level 1	Level 2	Level 3	Level 4	Unit
ΔSAT	Increase in mean annual global surface atmospheric temperature relative to 1850-1900	1.5	2	3	4	°C
MHW*	Global mean duration of marine heatwaves within a year	90	180	270	360	day
ΔSSL	Mean annual steric sea level rise relative to 1850-1900	0.2	0.3	0.4	0.5	m
SIE*	Arctic September sea-ice extent	4	3	2	1	10^6 km^2
ΔAMOC	Change in mean annual strength of the AMOC relative to 1850-1900	-20	-25	-30	-40	%
A$_{SO}$	Mean annual area proportion of Southern Ocean surface waters (south of 50°S) with aragonite undersaturation ($\Omega_{arag}<1$)	20	40	60	80	%

A_{Arctic}	Mean annual area proportion of Arctic Ocean surface waters (north of 70°N) with aragonite undersaturation ($\Omega_{arag}<1$)	20	40	60	80	%
$A_{\Omega<3}$	Mean annual area proportion of global ocean surface waters with $\Omega_{arag}<3$	50	70	90	100	%
Subsurface ΔO_2	Change in mean annual volume of hypoxic waters (<63 umol L^{-1}) above 1000 m relative to 1850-1900	2	4	6	8	%
Global ΔO_2	Change in mean annual global O_2 content relative to 1850-1900	-1.8	-2.4	-2.6	-3.5	%
ΔNPP^*	Change in mean annual depth-integrated net primary production relative to 1850-1900	-2	-3.5	-4	-8	%
$\Delta Biomass^*$	Change in mean annual depth-integrated plankton biomass relative to 1850-1900	-2	-3.5	-4	-8	%
$\Delta\Phi$	Change in mean annual upper-ocean (depth < 400 m) metabolic index relative to 1850-1900	-5	-10	-15	-20	%
ΔPOM	Change in mean annual particulate organic matter flux at 100 m depth (30°N-20°S) relative to 1850-1900	-4	-6	-8	-10	%

2.2 CMIP6 Earth system model ensemble

Our CMIP6 model ensemble is composed of 9 ESMs (Table 2). This ensemble is based on the one used in Canadell et al.

185 (2021), but excluding model family duplicates, and using the variant r1i1p1f1 (or equivalent). We use 3 scenarios from CMIP6 covering the period 2015-2100, which are initialized from the end of the historical simulation (1850 to 2014) that is based on estimates of historical forcings (O'Neill et al., 2016). These scenarios cover very different possible futures: The low-emission high-mitigation scenario SSP1-2.6 assumes that the world gradually shifts toward a more sustainable pathway, and that early and consistent climate mitigation limits the end-of-century radiative forcing to 2.6 W m^{-2}. In contrast, the SSP5-8.5 scenario

190 assumes resource-intensive, strong economic growth based on the exploitation of fossil fuel reserves and no climate mitigation. The very high CO_2 emissions in this scenario lead to a radiative forcing of 8.5 W m^{-2} at the end of this century. The SSP5-3.4-OS scenario follows the SSP5-8.5 pathway up to year 2040. Then, strong climate mitigation policies are implemented, including carbon dioxide removal from the atmosphere, leading to a peak and decline in surface temperature and a final

radiative forcing level of 3.4 W m^{-2} in 2100. To use the same model ensemble for all scenarios, we excluded models that do
195   not provide SSP5-3.4-OS.

Table 2. CMIP6 ensemble and variable availability per model.

Model	Reference	Variable availability
ACCESS-ESM1-5	Ziehn et al. (2020)	SAT, SIE, O_2, Φ, MHW, AMOC, Ω_{arag}, SSL, POM, NPP
CanESM5	Swart et al. (2019)	SAT, SIE, O_2, Φ, MHW, Biomass, AMOC, Ω_{arag}, SSL, POM
CESM2-WACCM	Danabasoglu et al. (2020)	SAT, SIE, Biomass, AMOC, Ω_{arag}, POM, NPP
CMCC-ESM2	Cherchi et al. (2019)	SAT, O_2, Φ, MHW, Biomass, Ω_{arag}, SSL, POM, NPP
CNRM-ESM2-1	Séférian et al. (2019)	SAT, SIE, O_2, Φ, MHW, Biomass, AMOC, Ω_{arag}, SSL, POM, NPP
IPSL-CM6A-LR	Boucher et al. (2020)	SAT, SIE, O_2, Φ, Biomass, AMOC, Ω_{arag}, SSL, POM, NPP
MIROC-ES2L	Hajima et al. (2020)	SAT, SIE, O_2, Φ, Biomass, AMOC, Ω_{arag}, POM, NPP
NorESM2-LM	Seland et al. (2020)	SAT, SIE, O_2, Φ, MHW, Biomass, AMOC, Ω_{arag}, SSL, POM, NPP
UKESM1-0-LL	Sellar et al. (2019)	SAT, SIE, O_2, Φ, MHW, Biomass, Ω_{arag}, SSL, POM, NPP

All ESM model outputs used in this study have been regridded to a 1°-resolution regular grid (360x180 grid cells) before
200   analysis. Since most of the impact metrics are expressed as a change relative to the period 1850-1900, model biases would
only be an issue for the analysis presented here if the response to forcing would significantly depend on the baseline state.
However, we removed potential model drifts from two sensitive metrics (ΔSSL and Global ΔO_2) by calculating the difference
between the projected signal and its equivalent from the corresponding preindustrial experiment (piControl). As previously
described, the impact metrics related to aragonite saturation state are also bias-corrected.

205   For all impact metrics, time series have been smoothed using a 20-year running mean before identifying the years and global
warming levels when a certain mitigation limit is exceeded. The exceedance is identified by the time and global warming level
at which a given mitigation limit is exceeded for the first time. This definition does not account for "overshooting" limits of
an impact metric. Cases where a limit is first exceeded, but the system returns to a state below the limit later in time, is counted
as an exceedance. Nevertheless, we provide analysis that allows for identifying cases where such overshooting of limit happens
210   (Figs. 3 and 4). Our analysis is complicated by the fact that some mitigation limits might be exceeded by only a part of all
available models or ensemble members, while other mitigation limits might be exceeded by all models or ensemble members
within the time horizon of the scenario simulations (until 2100). A model that does not exceed a given mitigation limit does
not provide any information on the time of exceedance. Consequently, our exceedance estimates are characterised in terms of
uncertainty and confidence. We define exceedance uncertainty as the interquartile range of an exceedance estimate in a given
215   model ensemble for a given experiment, impact metric, and limit. We define exceedance confidence as the proportion of
models exceeding a mitigation limit across all models that provide data for a given impact metric and limit. We assign high

confidence to an estimate of exceedance time (or warming level) if a mitigation limit is exceeded by at least 80 % of our CMIP6 ensemble (i.e., at least 8 out of 9 ESMs) or of the ensemble members of an individual EMIC, medium confidence if 50-79 % of the models or ensemble members exceed the limit, and low confidence if less than 50 % of models or ensemble members exceed the limit. One complication related to this methodology is that exceedance time (or warming levels) can be based on a different ensemble of ESMs for different scenarios, particularly if the exceedance estimates have different confidence levels. This could have been avoided by using ensemble means, and uncertainty intervals such as ensemble standard deviation to define exceedances, but a lot of information on the distribution of exceedances within the ensemble would have been lost leading to an underestimation of the resulting exceedance uncertainty. We will further discuss this issue below with some examples. We note that there are small differences as to which CMIP6 ESMs can be used for which metric, since not all models provide all data necessary for all impact metrics (Table 2).

2.3 Large perturbed-parameter ensembles from two Earth System Models of Intermediate Complexity

In addition to output from CMIP6 ESMs, we also analyse scenario simulations of two Earth system models of intermediate complexity (EMICs), the Bern3D-LPX model and the University of Victoria Earth System Climate Model (UVic). Both EMICs have simulated large perturbed-parameter ensembles to estimate the range of parametric (model) uncertainty. The EMIC perturbed-parameter ensembles were run over the historical period as well as for the SSP1-2.6, SSP5-3.4-OS, and SSP5-8.5 scenarios. Ensemble generation, sampled parameters, and calculation of ensemble member skill scores differ between the two models and are briefly outlined below.

Bern3D-LPX model

The model setup, ensemble generation, and evaluation as well as the experimental protocol of the Bern3D-LPX ensemble are the same as detailed in Jeltsch-Thömmes et al. (2024). The model features a three-dimensional dynamic ocean (Edwards et al., 1998; Müller et al., 2006) including sea-ice, a single-layer energy and moisture balance model of the atmosphere (Ritz et al., 2011), and a comprehensive terrestrial biosphere component (LPX-Bern v1.5) with dynamic vegetation, fire, nitrogen, nitrous oxide, methane, permafrost, peatland, and land-use modules (Lienert and Joos, 2018).

The sampling approach for the perturbed-parameter ensemble builds upon work by Steinacher et al. (2013) and is used thereafter in several follow-up studies (e.g., Steinacher and Joos, 2016; Battaglia et al., 2016; Battaglia and Joos, 2018; Lienert and Joos, 2018). A 1000-member perturbed parameter ensemble is generated from the prior distributions of 27 key model parameters using Latin hypercube sampling (Mckay et al., 2000; Steinacher et al., 2013).

To reduce uncertainties, we exploit a broad set of observation-based data (Fig. A1) to constrain the model ensemble to realisations that are compatible with observations, thereby probing both the mean state and the transient response in space and time of the ensemble members. Further details on the methods used to constrain the Bern3D-LPX model ensemble with observations are provided in Jeltsch-Thömmes et al. (2024).

250 **UVic ESCCEM v2.10**

The UVic ESCM v2.10 (Mengis et al., 2020; Weaver et al., 2001) has a three dimensional ocean with a horizontal resolution of 3.6° longitude, 1.8° latitude, and 19 vertical levels. The atmosphere is a two-dimensional energy-moisture balance model with the same horizontal resolution. The oceanic physics follows the Modular Ocean Model version 2 (MOM2) (Pacanowski, 1996) and the ocean biogeochemistry model is outlined by (Keller et al., 2012). A thermodynamic-dynamic sea ice model

255 (Bitz et al., 2001) employing elastic visco-plastic rheology (Hunke and Dukowicz, 1997) is coupled to the ocean. The atmosphere is represented by a two-dimensional atmospheric energy moisture balance model (Fanning and Weaver, 1996). The terrestrial component accounts for vegetation dynamics and incorporates five different plant functional types (Meissner et al., 2003). Additionally, the model includes a representation of permafrost carbon (MacDougall et al., 2017) using a diffusion-based scheme, which approximates the process of cryoturbation.

260 A 325-member perturbed parameter ensemble is generated using a multi-wave history matching approach (Andrianakis et al., 2015; Bower et al., 2010). History matching (HM), or iterative refocusing, is based on running an ensemble in a predefined parameter space, using it to train statistical emulators that predict key metrics from the model output, and then using the emulator to identify the set of inputs that would give an acceptable match between the model output and the observed data. In our case, we performed six waves of 80 simulations each and compared model outputs with observations after each wave.

265 Gaussian Process (GP) emulators (Kennedy and O'Hagan, 2001; Rasmussen and Williams, 2005; Sacks et al., 1989) are then constructed to predict these outputs as functions of the perturbed parameters to reject regions of the input space which are unlikely to produce results consistent with observations. For each quantity that we compare to observation, an implausibility measure (Andrianakis et al., 2015; Williamson et al., 2015) is computed following Eq. (2):

$$I_j(x) = \frac{|z_j - E^*[g_j(x)]|}{[V_o + V_c(x) + V_m]^{1/2}}, \tag{2}$$

270 where $g_j(x)$ is the function describing the relationship between a vector of model inputs x and a specific model output j. Since we employ GP emulation, we have the expectation provided by the emulators $E^*[g_j(x)]$. The corresponding observation is z_j. The term V_o, $V_c(x)$, and V_m represent the variance associated with the observational uncertainty, the code uncertainty as given by the emulator, and the model discrepancy. The latter is simply defined as 10 % of the ensemble range due to the difficulty to estimate model discrepancy. The value of I is large if it is unlikely for the model to produce an acceptable match with

275 observation when using the input combination x. We adopt a similar approach as described in Jeltsch-Thömmes et al. (2024) for the Bern3D-LPX model to compute a score S based on our calculated implausibility measure $I_j(x)$. We generate a large Latin hypercube sampling plan and reject parameter combinations with emulated $I_j(x) > 3$. The emulated 1978-member ensemble was weighted using the score S and used for all statistical computations in this work.

3 Results and Discussion

280 **3.1 CMIP6 ESM**

The uncertainty, and confidence related to the time and global warming levels at which mitigation limits are exceeded are highly variable across impact metrics, limits, and scenarios (Figs. 1 and 2). The confidence in exceedance estimates decreases with higher mitigation limits (since generally less models exceed the higher mitigation limits), except for A_{SO} and SIE. Almost all (except for ΔSSL and A_{SO}) of the most ambitious mitigation limits (limit 1) are exceeded in the short- to mid-term (before

285 2060) for all scenarios. If mitigation limit 1 is exceeded with high confidence (marked by full size black dots in Figs. 1 and 2) in all scenarios, the median time of exceedance is generally very similar across scenarios (ΔSAT, SIE, $\Delta Biomass$, A_{Arctic}), because during earlier times the three scenarios share the same historical forcing or have only slightly diverged. Also, before 2040, the two scenarios of the SSP5 family are identical by construction, such that if all models exceed a mitigation limit before 2040, the median exceedance time is identical for SSP5-3.4-OS and SSP5-8.5. Lower exceedance confidence and larger

290 difference in the timing of exceedance are found for ΔPOM, ΔNPP, global ocean $\Delta O2$, and subsurface $\Delta O2$. The low confidence in exceedance of ΔNPP mitigation limits across all scenarios is consistent with the high uncertainty in ΔNPP projections found by Kwiatkowski et al. (2020), while, in contrast, the higher confidence for the metric $\Delta Biomass$ is consistent with the findings of Tittensor et al. (2021). Substantial uncertainties in ΔO_2 were found in earlier multi-model studies (Cocco et al., 2013; Hameau et al., 2020).

295

The varying exceedance confidence of some metrics and limits complicates the comparison across scenarios because of the difference in the number models contributing to the exceedance distribution. This can lead to counterintuitive results. For example, in Fig. 1, for the first limit of ΔPOM, the 25th percentile and the median of the exceedance years varies across scenarios before 2015 even though the historical period (1850-2014) is identical for all scenarios. On another hand, the

300 exceedance distribution of the first limit of $\Delta Biomass$ shows earlier median exceedances under the low-emissions SSP1-2.6 scenario because one model does not exceed the limit under SSP1-2.6 at all (MIROC-ES2L, not shown), while this model exceeds the same limit the latest under the high-emissions SSP5-8.5 scenario (year 2053).

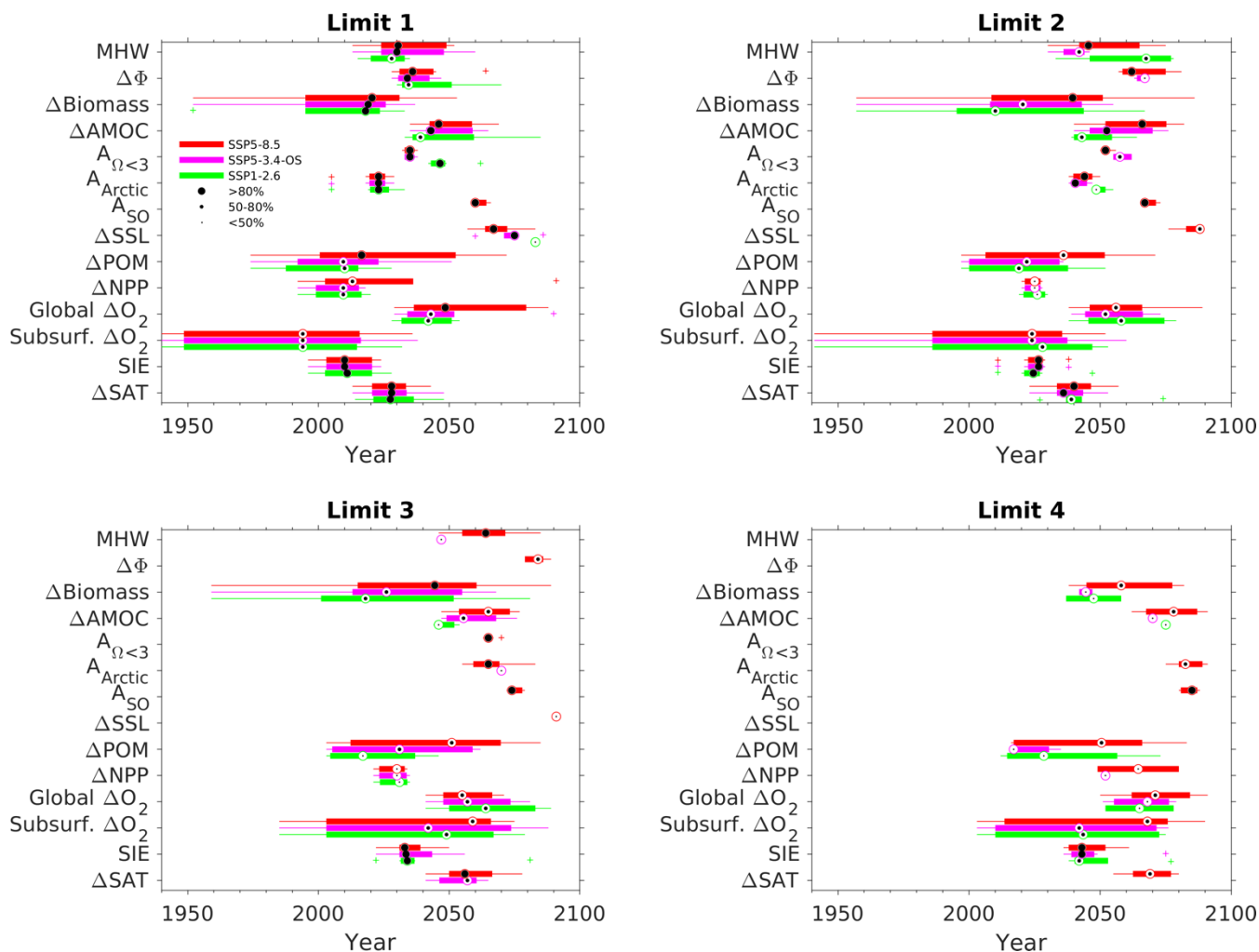



Figure 1: Box plots showing the distribution of exceedance years for each mitigation limits of the impact metrics (abbreviations follow Table 1) for CMIP6 models. Green, purple and red colors depict the three scenarios historical-SSP1-2.6, historical–SSP5-3.4-OS and historical–SSP5-8.5. Boxes' lengths and circled black dots depict the interquartile range and the ensemble median, respectively. Whiskers extend to the most extreme data value that is not considered as an outlier (i.e., not greater than 1.5 times the interquartile range). Crosses indicate outliers. The size of the black dots indicates the exceedance confidence defined as the proportion of models exceeding a mitigation limit across the models providing data for a given impact metric (Full size: >80 %, half size: 50-80 %, small size: <50 %).

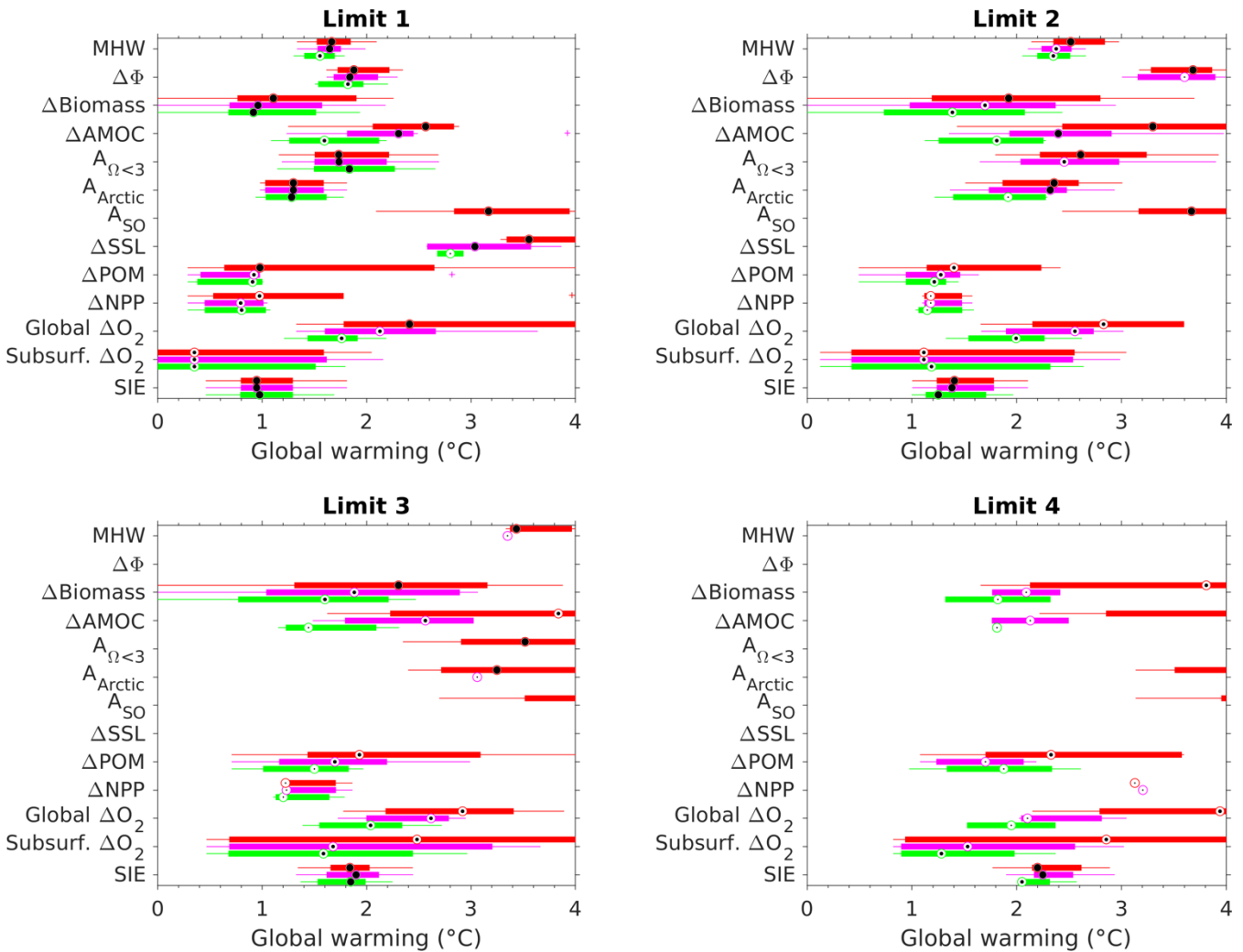



Figure 2: Same as Figure 1, but relative to global warming during this century instead of time.

315   For the less ambitious mitigation limits, exceedance time estimates generally move towards later times and higher warming levels, and the exceedance confidence decreases (less models exceed the less ambitious mitigation limits). In the low emission scenario SSP1-2.6, limit 4 is not exceeded by more than 80 % of the CMIP6 ESMs for any of the impact metrics due to the low radiative forcing in SSP1-2.6 compared to, for example, SSP5-8.5 scenario. However, there are some metrics not following this behaviour: In Figure 2, the exceedance distribution of the fourth limit of subsurface ΔO_2 under SSP1-2.6 is centred on

320   lower global warming levels than the one of the third limit (Figure 2). This shift in the distribution is due to some models exceeding only the third limit, and consequently the exceedance confidence is lower for the fourth limit. Similarly, many exceedances occur at lower global warming level under SSP1-2.6 scenario compared to SSP5-8.5 scenario (e.g., in limit 2, ΔBiomass, ΔAMOC, ΔPOM, and global ΔO_2). This is explained by the sustained but slow-paced changes of these metrics that eventually leads to an exceedance even under the global warming stabilization induced by SSP1-2.6 scenario.

325   If we focus on exceedance estimates that have a high confidence (i.e., where more than 80 % of models exceed the mitigation limit), we can provide an estimate of the time when mitigation limits are likely to be exceeded (Figs. 3 to 6). The first mitigation limit of A_{Arctic} is likely already passed in the CMIP6 model ensemble in all scenarios, consistent with the findings of Terhaar et al. (2021). The first mitigation limit of SIE is expected to be passed during 2013-2034 (median year 2023). So far, according to satellite-based estimates, Arctic summer sea-ice extent fell below $4 \cdot 10^6$ km^2 only in 2012 and 2020 during the last decade

330   (https://nsidc.org/, visited on March 14th, 2024). The fourth level of mitigation limits is exceeded with high confidence only for A_{SO} in SSP5-3.4-OS and SSP5-8.5 as well as for SIE in SSP5-8.5.

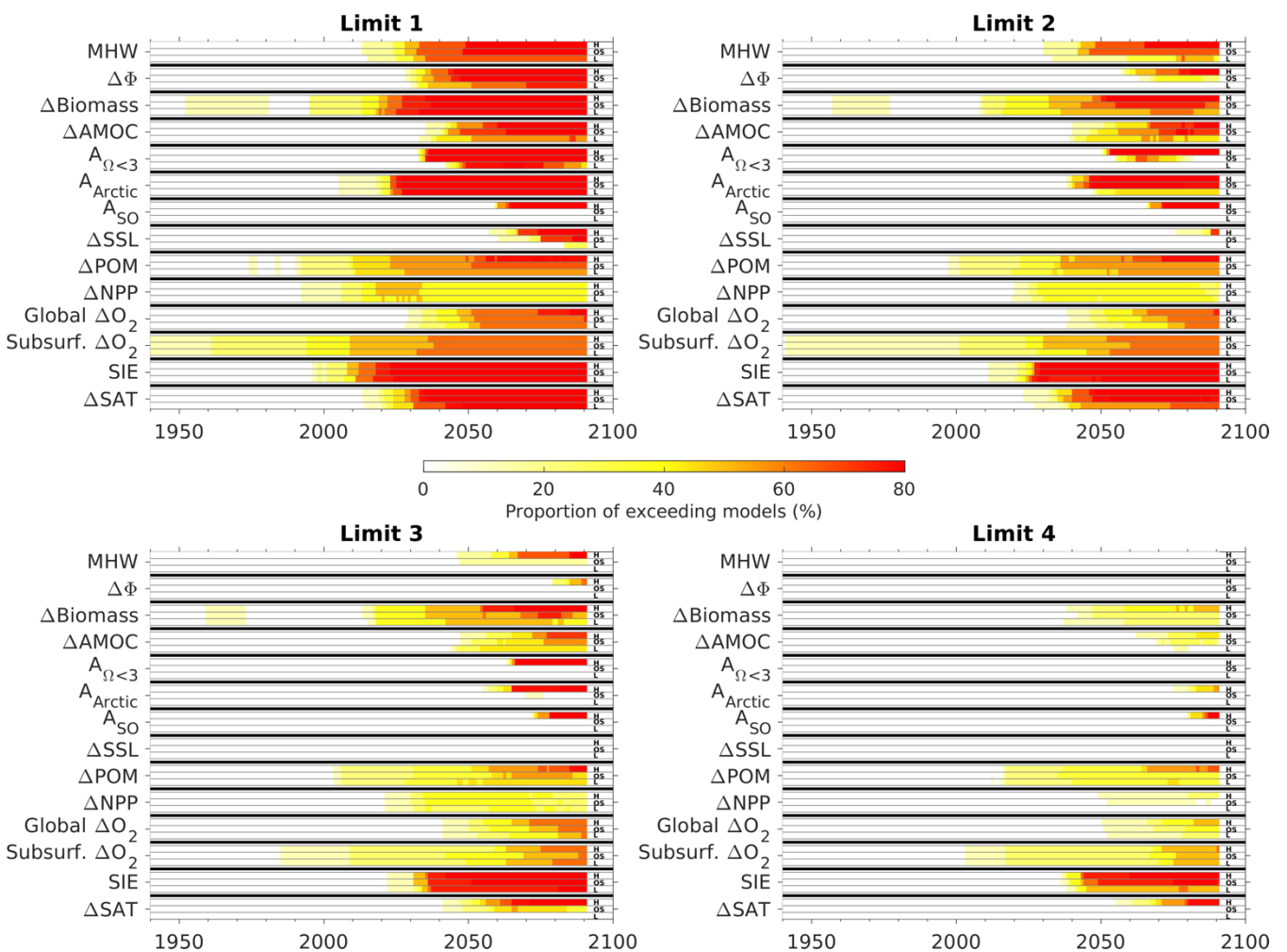



Figure 3: Proportion of models exceeding a given mitigation limit of the impact metrics compared to the available models for each
335   **metric. Abbreviations are included for SSP1-2.6 (L), SSP5-3.4-OS (OS) and SSP5-8.5 (H).**

Avoiding emissions as high as in the SSP5-8.5 scenario and following an emission pathway similar to SSP5-3.4-OS will likely avoid an exceedance of any of the mitigation limits for A_{SO}, ΔPOM, and Global ΔO_2 during this century. In addition, avoiding

the SSP5-3.4-OS scenario by early mitigation (as in SSP1-2.6) will likely avoid exceedances of any mitigation limit related to MHW, ΔBiomass, ΔΦ and ΔAMOC until year 2100. The mitigation limits of $A_{\Omega<3}$, A_{Arctic}, subsurface ΔO_2, SIE, ΔSSL and

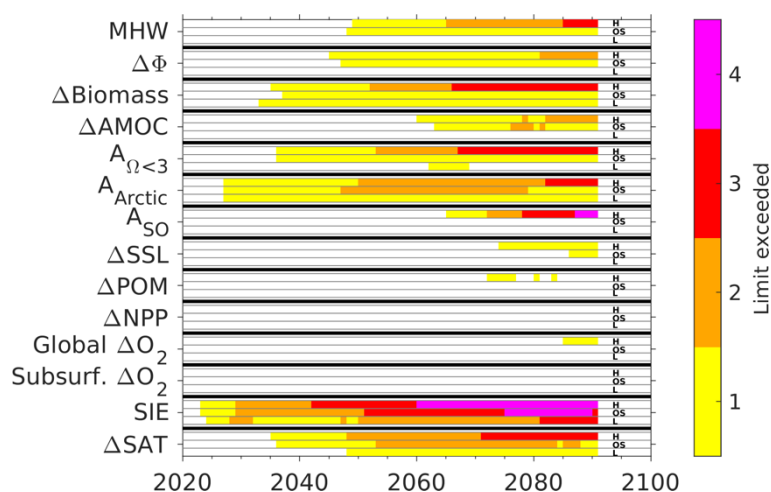
340 ΔSAT are likely to be exceeded across all scenarios. The effect of ambitious mitigation in SSP5-3.4-OS after 2040 can be seen for ΔSAT, SIE, ΔAMOC, ΔSSL and ΔBiomass. For these metrics, some of the mitigation limits are first exceeded around mid-century, but later this exceedance is reversed. The uncertainty of ΔNPP, ΔPOM, global and subsurface ΔO_2 projections does not allow us to conclude with confidence that none of the corresponding mitigation limits will be exceeded (Fig. 4).

345

Figure 4: Time periods where mitigation limits are exceeded with high confidence according to the CMIP6 model ensemble (>80 % of models) for each impact metrics and for the scenarios SSP1-2.6 (L), SSP5-3.4-OS (OS) and SSP5-8.5 (H).

In Figs. 5 and 6, the summary of high confidence exceedance estimates for all impact metrics is quite conservative, since (1) high confidence is defined as at least 80 % of models exceeding a limit (i.e., at least 8 out 9 models, which is practically 88

350 %) and (2) medium confidence exceedances (where up to 79 % of the CMIP6 models would show an exceedance of a given mitigation limit) are not included. For the low-emission scenario, already the two most ambitious levels of limits of a few impact metrics (ΔSAT, SIE, A_{Arctic}, and $A_{\Omega<3}$) are exceeded with high confidence. In contrast, most of the mitigation limits are exceeded with high confidence in the high-emission SSP5-8.5 scenario, particularly toward 2100. The absence of high confidence in the exceedance for global and subsurface ΔO_2 is explained partly by a model disagreement within our CMIP6

355 ensemble. Another reason for low-to-medium confidence in the exceedance for global and subsurface ΔO_2 and ΔSSL before year 2100 is that changes in subsurface and whole ocean parameters have been shown to accrue beyond year 2100 and aggravate over many centuries due to the long overturning time scales of the ocean (Battaglia and Joos, 2018). There is a very clear effect of the ambitious mitigation assumed in SSP5-3.4-OS after 2040 in all timeseries of impact metrics, such that the significantly lower exceedance rate of mitigation limits, particularly by the end of the century, clearly illustrated the benefits

360 of stringent and ambitious mitigation. Some metrics show strong hysteresis under cumulative carbon emissions (Boucher et al., 2012; Jeltsch-Thömmes et al., 2020; Samanta et al., 2010; Santana-Falcón et al., 2023). Due to hysteresis, sustained

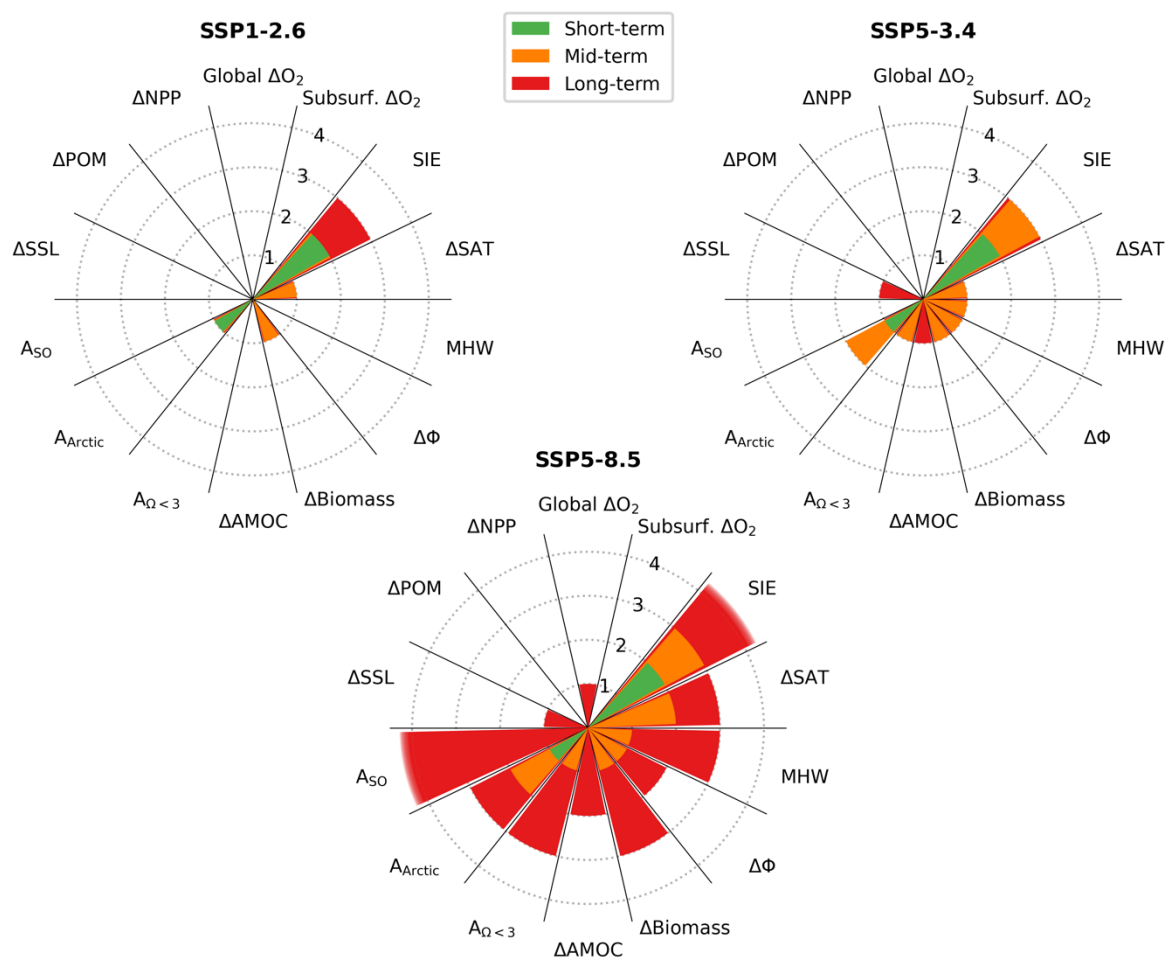
negative emissions are required to return to and stay under a specific limit, particularly for high climate sensitivities and peak-and-decline scenarios with carbon dioxide removal (Jeltsch-Thömmes et al., 2020). This aspect needs to be emphasized in the case of our study due to the use of simulations ending in 2100.

365

Figure 5: Exceedance of mitigation limits with high confidence (>80 % of the CMIP6 models) in the near-term (2021-2040), mid-term (2041-2060) and long-term (2081-2100) periods under (left) SSP1-2.6, (right) SSP5-3.4-OS and (middle) SSP5-8.5 scenarios.

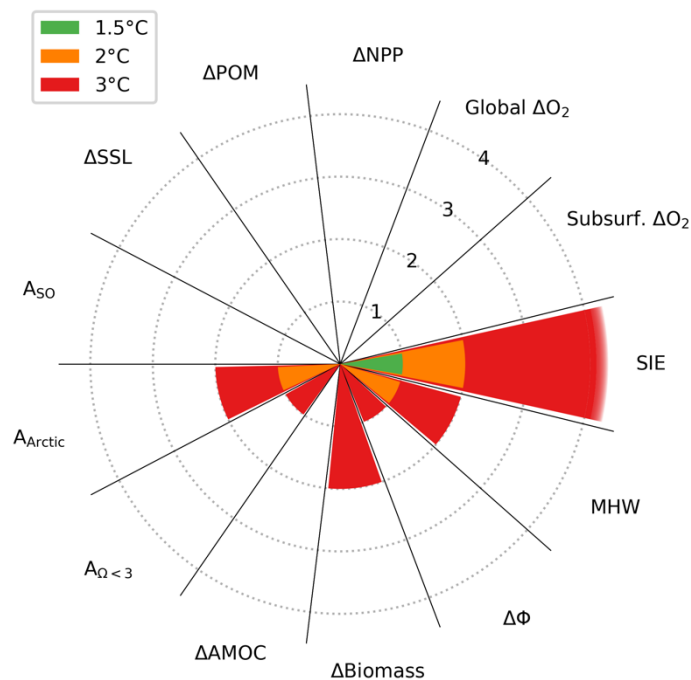



Figure 6: Same as Figure 5 according to global warming levels (1.5°C, 2°C and 3°C) under SSP5-8.5 scenario.

370 **3.2 EMIC & ESM comparison**

There is broad agreement for a range of variables on the limit exceedances between the CMIP6 and the two (skill-weighted) EMIC ensembles under SSP5-8.5, but also major disagreements are identified (Figs. 7 and 8). Note that the median and interquartile ranges are only given for those models or ensemble members that pass the limits. Thus, for a good agreement between ensembles both confidence and median needs to match. For the limit set 2, the median value of exceedance agrees

375 within 25 years and 1°C for ΔAMOC, $A_{\Omega<3}$, A_{Arctic}, A_{SO}, ΔPOM, and ΔSAT between the CMIP6 and the two EMIC ensembles. However, confidence is variable among ensembles. The two EMIC ensembles generally show systematically high confidence in limit exceedances in the first two limit sets (except UVic's $\Delta\Phi$ limit 2) while the CMIP6 ensemble shows 5 exceedances with medium confidence over the same limit sets. The CMIP6 ensemble shows somewhat larger warming than the EMIC ensemble with earlier exceedance but a smaller fraction of CMIP6 models exceeding the limits than the (skill-score weighted)

380 EMIC ensembles. This is consistent with the fact that the CMIP6 ensemble includes models with climate sensitivity larger than observation-constrained estimates (Nijsse et al., 2020; Tokarska et al., 2020). The CMIP6 models show medium confidence in the exceedance of the subsurface O_2 for the third and fourth limit sets, whereas the EMIC ensembles show no exceedance or with little confidence. The finer spatial resolution used in CMIP6 models compared to the EMIC ensemble could explain this difference, particularly for subsurface O_2. The Bern3D-LPX model is the only ensemble showing high

385 confidence in all exceedances.

Regarding to the exceedance uncertainties represented by interquartile ranges in Fig. 7, the three ensemble agrees generally well over many metrics, except a few metrics such as ΔPOM and subsurface ΔO_2. For these metrics, the uncertainty range from the CMIP6 ensemble is larger than the EMIC ensembles. We hypothesize that the parametric uncertainty sampled in the perturbed-parameter EMIC ensembles is a significant underestimation of the full uncertainty signal of some metrics as it lacks

390   the structural model uncertainty inherent to the CMIP6 ensemble.

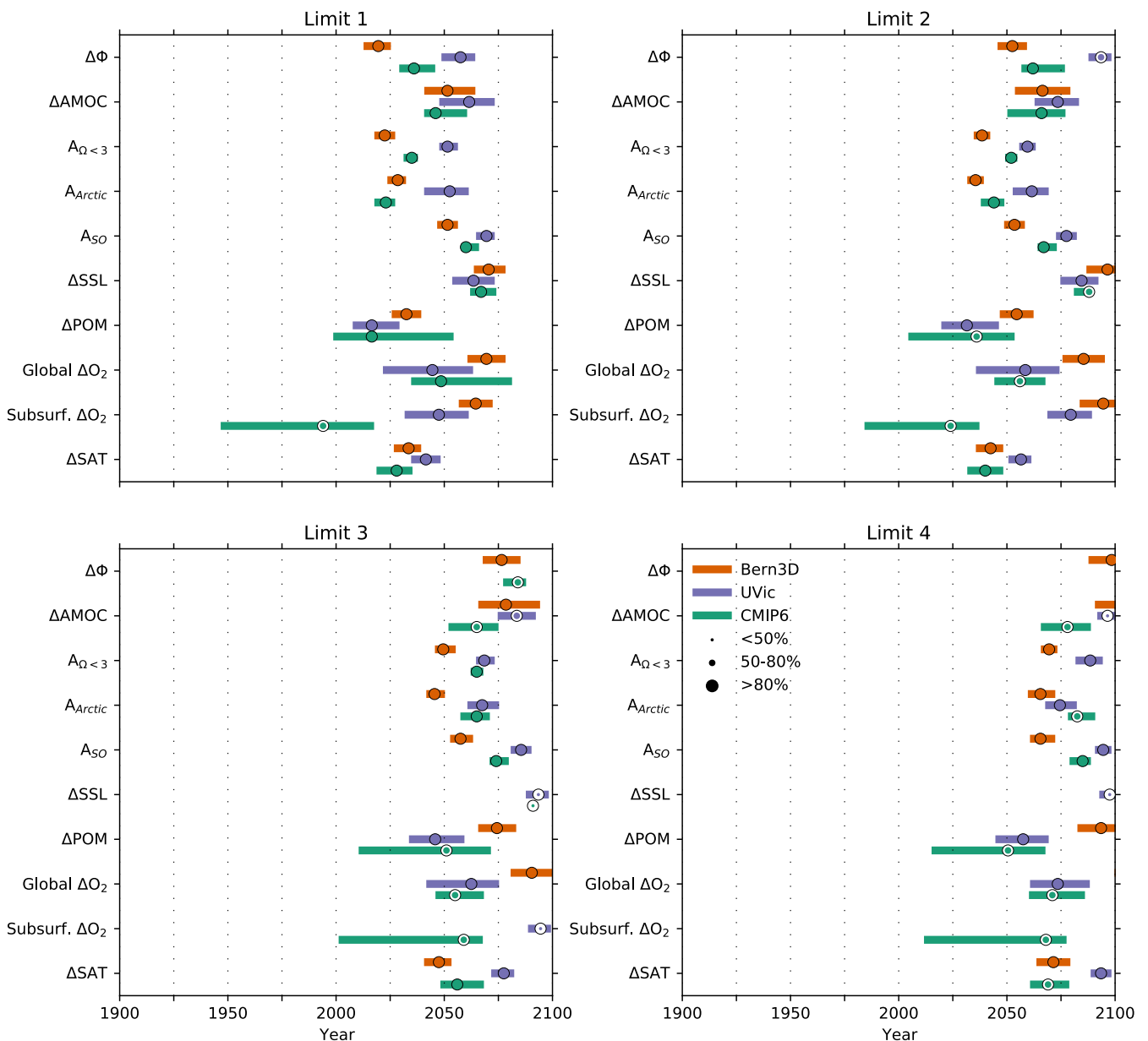



Figure 7: Box plots showing the interquartile range distribution of exceedance years for each mitigation limit of the impact metrics (abbreviations follow Table 1) for SSP5-8.5. Orange, purple, and green show data from the Bern3D-LPX, UVic, and CMIP6

395 **ensemble, respectively. Dots indicate the median and the size of the dots indicates the percentage of ensemble members that have crossed the respective limit. Note that the median and interquartile ranges are only given for those models or ensemble members that pass the limits. MHW, SIE, NPP and ΔBiomass are not shown because EMIC ensembles were not able to provide data for these metrics.**

Despite such differences, robust conclusions emerge. First, both the CMIP6 and EMIC ensembles demonstrate that most of the stringent limits of set 1 and set 2 are passed with high confidence within this century for global warming of 2°C and 2.5°C,

400 respectively (Fig. 8). Exceptions are $\Delta\Phi$, ΔSSL (known to lag surface warming and continues to increase over centuries), and subsurface ΔO_2, for which the CMIP6 and EMIC ensembles disagree. Second, both the CMIP6 and EMIC ensembles demonstrate that many less stringent limits of set 3 and set 4 are not passed with high confidence within this century for global warming of 1.5°C and 2°C, respectively (Fig. 8). Taken together, the results of the model ensembles collectively show that limiting global warming below 2°C avoids passing the considered Earth system limits during this century with potentially

405 dangerous impacts on eco- and socio-economic systems.

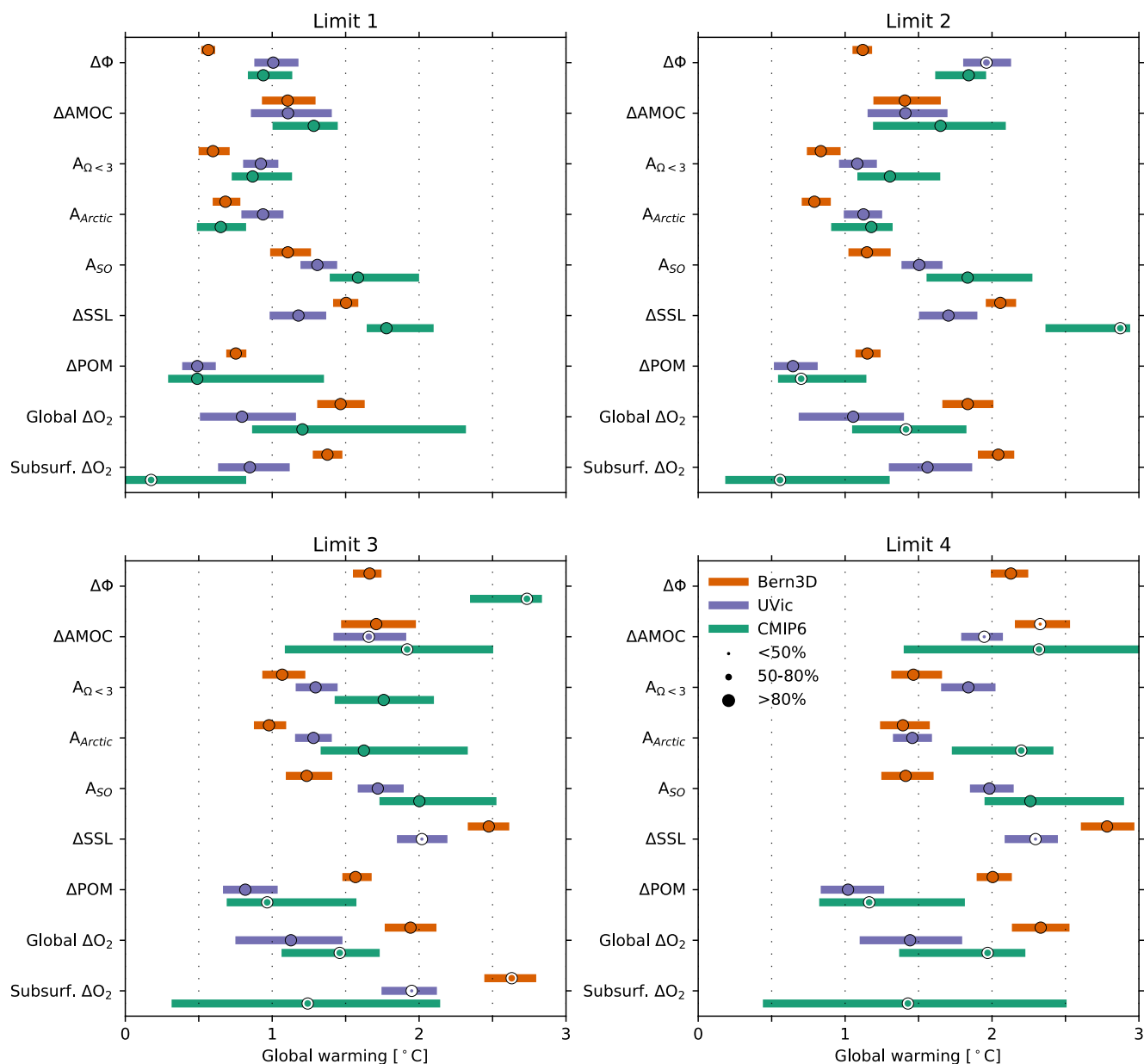



Figure 8: Same as Figure 7 but according to global warming following ΔSAT definition.

Conclusion

This study assesses different types of IPCC emission pathways (low, high, overshoot) with respect to multi-dimensional safe marine operating spaces including a wide range of ocean impact metrics and corresponding mitigation limits based on the literature. It contributes to identifying viable mitigation pathways for the 21st century projected by state-of-the-art Earth System

Models, complemented by two observation-constrained ensembles from Earth System Models of Intermediate Complexity. Assessing the exceedance of mitigation limits for multiple impact metrics requires large model ensembles to be able to obtain high-confidence signals and corresponding uncertainties for the exceedance estimates (years and global warming levels) linked

415 to the projection pathways. The large uncertainties found for exceedance estimates of many of the impact metrics highlight the need for better constraining and/or weighting the CMIP6 ensemble. Simulations beyond year 2100 are needed to assess the long-term impacts of anthropogenic emissions, especially for subsurface and global oxygen, and steric sea level rise.

Our results show that ambitious mitigation limits will be exceeded with high or medium confidence even if a low-emission pathway is followed, but that exceeding less ambitious mitigation limits (associated with a higher risk for severe impacts) is

420 unlikely in a low-emission scenario. In contrast, under the high-emission scenario, many of the less ambitious and more risk-prone mitigation limits are exceeded with high to medium confidence. The benefit of strong mitigation efforts in the overshoot pathway is clearly measurable as a decrease in the exceedance probability of the least ambitious and most risk-prone mitigation limits.

Appendix

425

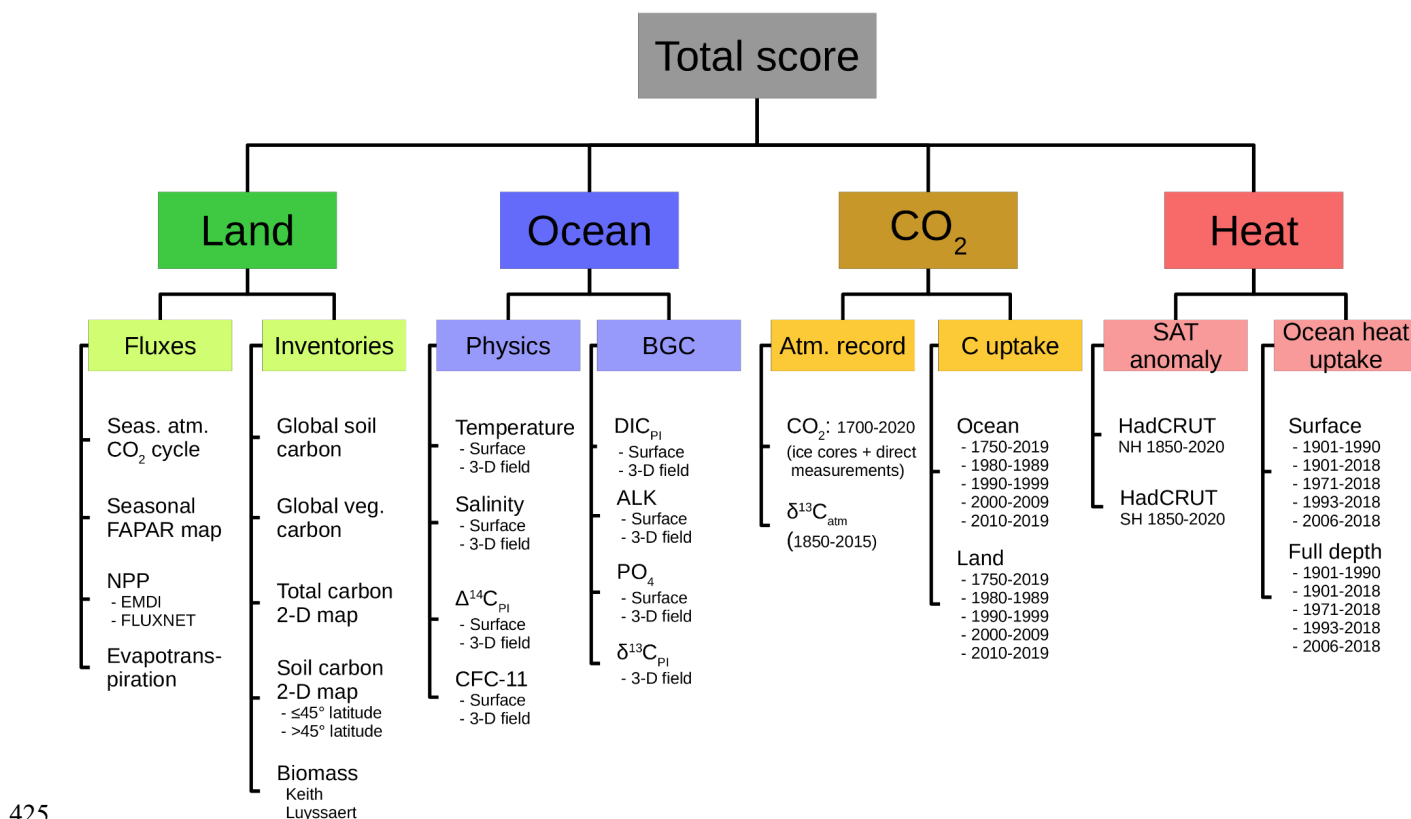



Figure A1: Hierarchical weighting scheme used to calculate the skill scores of individual ensemble members of the Bern3D-LPX model ensemble. Data sets at each level have equal weight. For example, the data-model mismatch in "Surface DIC" in the entry

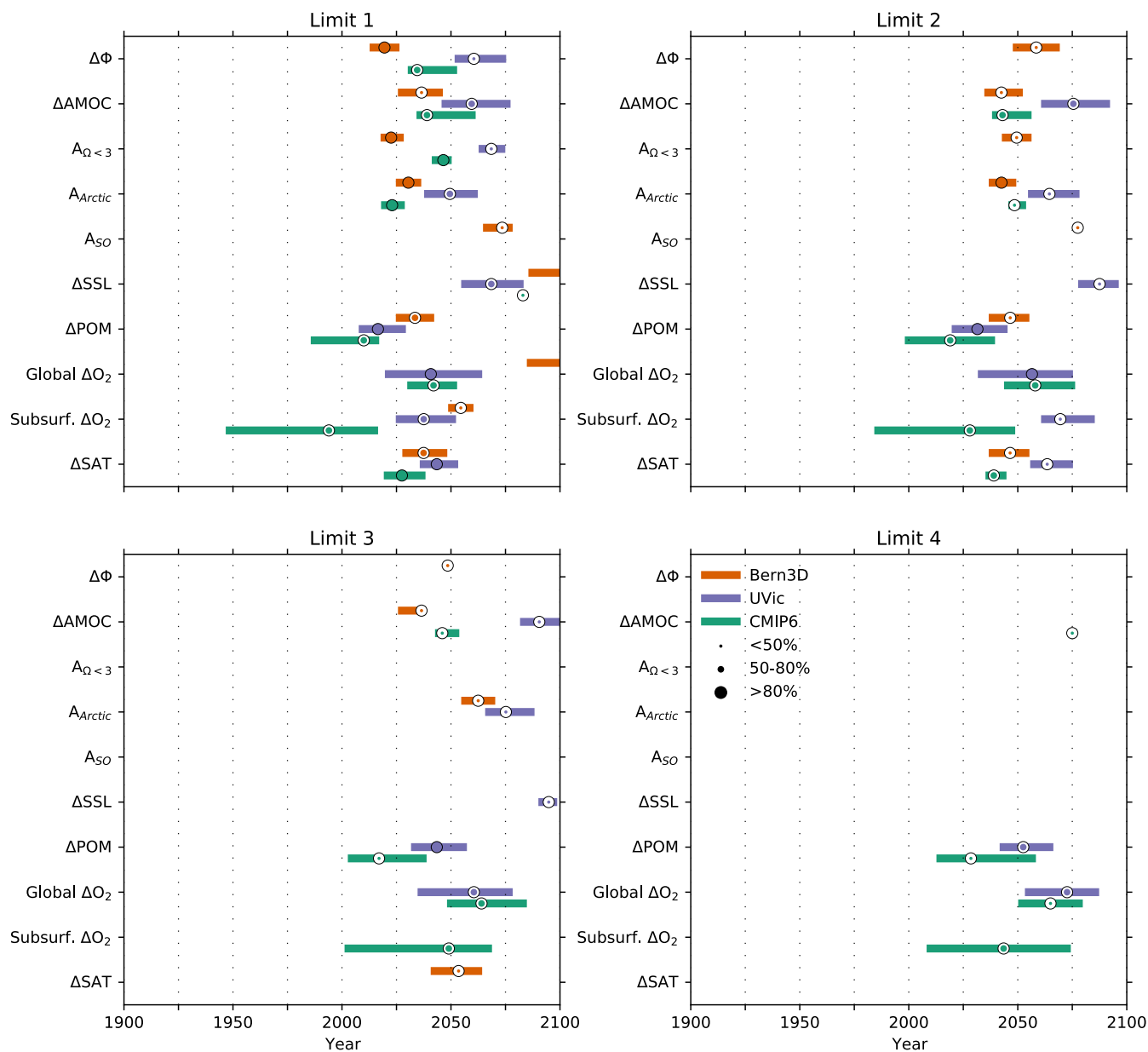
Ocean - BGC enters the total skill with a weight of 1/64 (½ (two data sets: Surface and 3-D field) x ¼ (4 groups: DIC, ALK, PO4, 13C) x ½ (two major subgroups: Physics, BGC) x ¼ (4 major groups: Land, Ocean, CO2, Heat)).

430

Figure A2: Same as in Figure 7 but for SSP1-2.6.

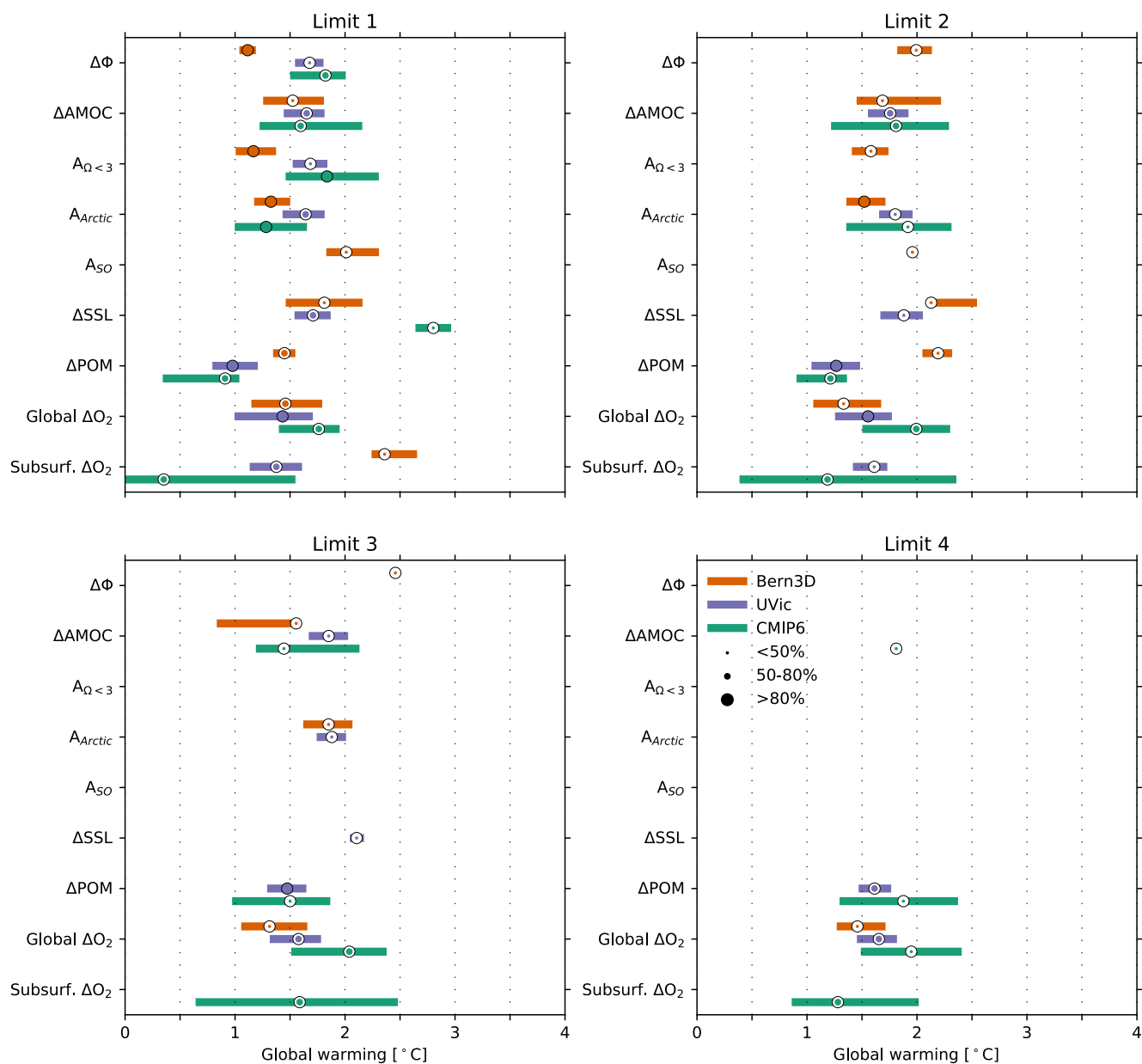
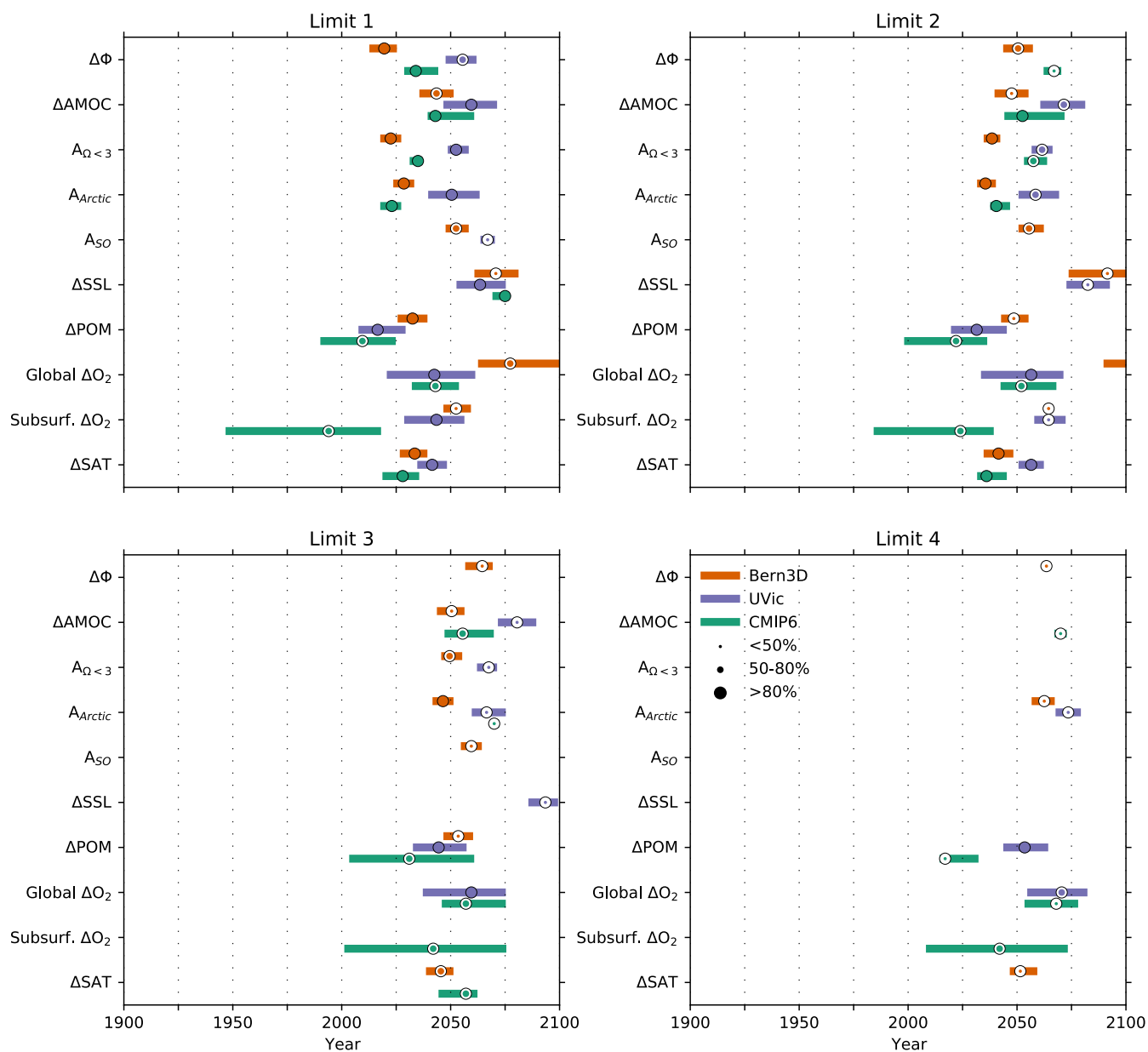



Figure A3: Same as in Figure 8 but for SSP1-2.6.

435 **Figure A4: Same as in Figure 7 but for SSP5-3.4-OS.**

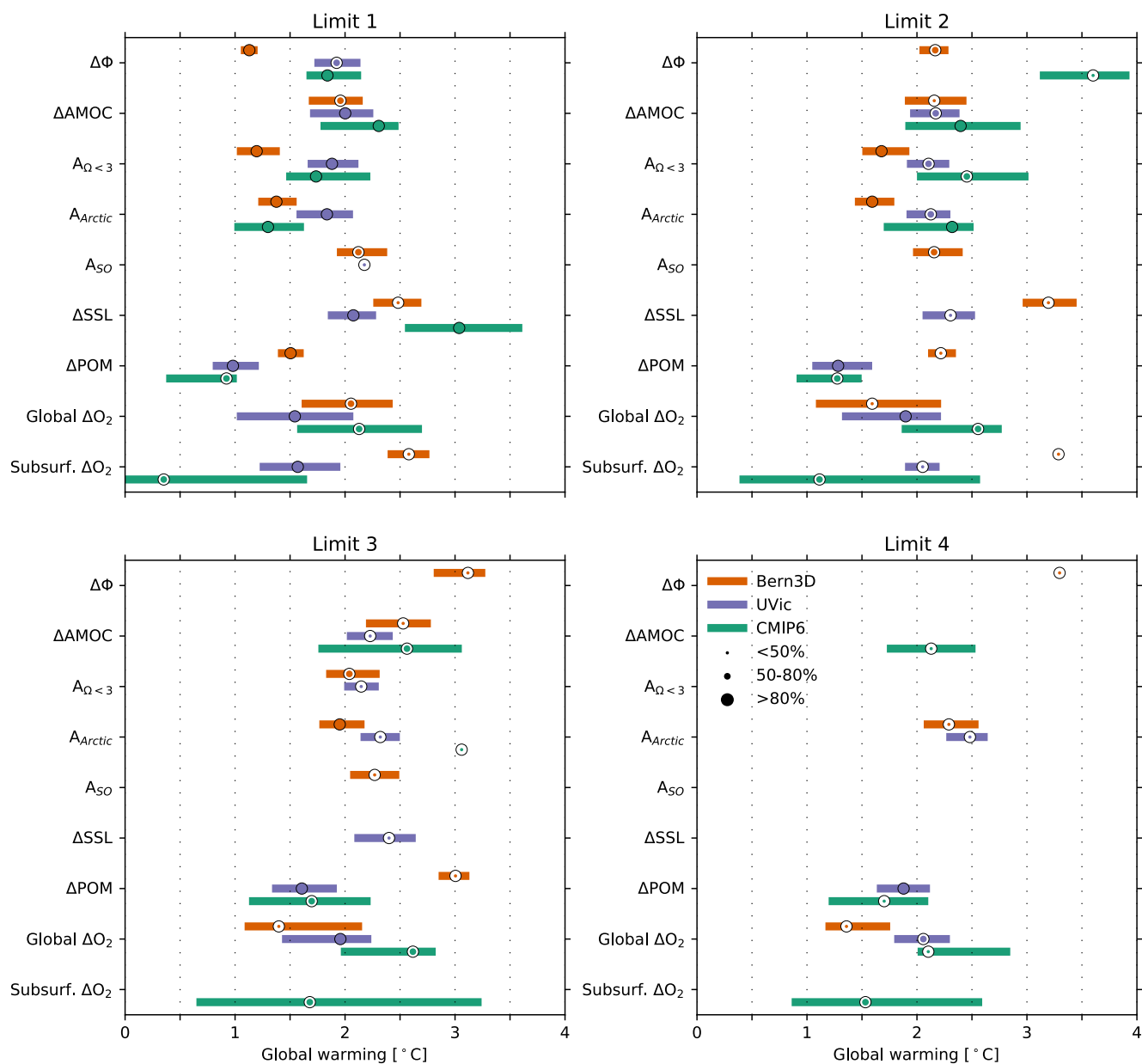



Figure A5: Same as in Figure 8 but for SSP5-3.4-OS.

Code availability

The mocsy 2.0 code is publicly available via https://github.com/jamesorr/mocsy (Orr and Epitalon, 2015).

Data availability

CMIP6 outputs are publicly available from the Earth System Grid Federation (ESGF) portals (e.g. https://esgf-data.dkrz.de/).
The World Ocean Atlas 2013 (Locarnini et al., 2013; Zweng et al., 2013; https://www.nodc.noaa.gov/OC5/woa13/) and the
GLODAPv2 (Lauvset et al., 2016; https://www.nodc.noaa.gov/ocads/oceans/GLODAPv2_2019/) data products are available
from the National Oceanographic Data Center portal of the National Oceanic and Atmospheric Administration.

Author contribution

The study was led by T.B. who performed the CMIP6 calculations, analysis, and figures. O.T. processed the CMIP6 model
data (download, regridding). F.F. provided the CMIP6 metabolic index data. A. J.-T. analysed the Bern-3D data and provided
the EMICs figures. G. T. T. led the metrics' and limits' literature review and analysed the UVic data. T. L. F. provided the
CMIP6 marine heatwave data. J. N. provided the summary Figs. 5 and 6. All authors were involved in designing the analysis,
interpreting the results and writing the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Disclaimer

The work reflects only the authors' view; the European Commission and their executive agency are not responsible for any
use that may be made of the information the work contains.

Acknowledgments

We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling,
coordinated and promoted CMIP. We thank the climate modelling groups for producing and making available their model
output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding
agencies who support CMIP and ESGF. All authors received funding from the European Union's Horizon 2020 research and
innovation programme under grant agreement No. 820989 (COMFORT). FJ and AJ acknowledge additional funding by the
Swiss National Science Foundation (#200020_200511). TLF received also funding from the European Union's Horizon
Europe research and innovation programme under grant agreement No. 101137673 (TipESM). We also thank the IPSL
modelling group for the software infrastructure, which facilitated CMIP6 analysis.

References

Anderson, S. I., Barton, A. D., Clayton, S., Dutkiewicz, S., and Rynearson, T. A.: Marine phytoplankton functional types exhibit diverse responses to thermal change, Nat Commun, 12, 6413, https://doi.org/10.1038/s41467-021-26651-8, 2021.

Andrianakis, I., Vernon, I. R., McCreesh, N., McKinley, T. J., Oakley, J. E., Nsubuga, R. N., Goldstein, M., and White, R. G.: Bayesian History Matching of Complex Infectious Disease Models Using Emulation: A Tutorial and a Case Study on HIV in Uganda, PLOS Computational Biology, 11, e1003968, https://doi.org/10.1371/journal.pcbi.1003968, 2015.

Armstrong McKay, D. I., Staal, A., Abrams, J. F., Winkelmann, R., Sakschewski, B., Loriani, S., Fetzer, I., Cornell, S. E., Rockström, J., and Lenton, T. M.: Exceeding 1.5°C global warming could trigger multiple climate tipping points, Science, 377, eabn7950, https://doi.org/10.1126/science.abn7950, 2022.

Battaglia, G. and Joos, F.: Hazards of decreasing marine oxygen: the near-term and millennial-scale benefits of meeting the Paris climate targets, Earth System Dynamics, 9, 797–816, https://doi.org/10.5194/esd-9-797-2018, 2018.

Battaglia, G., Steinacher, M., and Joos, F.: A probabilistic assessment of calcium carbonate export and dissolution in the modern ocean, Biogeosciences, 13, 2823–2848, https://doi.org/10.5194/bg-13-2823-2016, 2016.

Bitz, C. M., Holland, M. M., Weaver, A. J., and Eby, M.: Simulating the ice-thickness distribution in a coupled climate model, Journal of Geophysical Research: Oceans, 106, 2441–2463, https://doi.org/10.1029/1999JC000113, 2001.

Bopp, L., Resplandy, L., Orr, J. C., Doney, S. C., Dunne, J. P., Gehlen, M., Halloran, P., Heinze, C., Ilyina, T., Séférian, R., Tjiputra, J., and Vichi, M.: 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, 2013.

Bopp, L., Aumont, O., Kwiatkowski, L., Clerc, C., Dupont, L., Ethé, C., Gorgues, T., Séférian, R., and Tagliabue, A.: Diazotrophy as a key driver of the response of marine net primary productivity to climate change, Biogeosciences, 19, 4267–4285, https://doi.org/10.5194/bg-19-4267-2022, 2022.

Boucher, O., Halloran, P. R., Burke, E. J., Doutriaux-Boucher, M., Jones, C. D., Lowe, J., Ringer, M. A., Robertson, E., and Wu, P.: Reversibility in an Earth System model in response to CO_2 concentration changes, Environ. Res. Lett., 7, 024013, https://doi.org/10.1088/1748-9326/7/2/024013, 2012.

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D'Andrea, F., Davini, P., de Lavergne, C., Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J., Guenet, B., Guez, L., E., Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T., Luyssaert, S., Madec, G., Madeleine, J., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R., Traore, A. K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N.: Presentation and Evaluation of the IPSL-CM6A-LR Climate Model, J Adv Model Earth Syst, 12, https://doi.org/10.1029/2019MS002010, 2020.

Bower, R. G., Goldstein, M., and Vernon, I.: Galaxy formation: a Bayesian uncertainty analysis, Bayesian Analysis, 5, 619–669, https://doi.org/10.1214/10-BA524, 2010.

Canadell, J. G., Monteiro, P. M. S., Costa, M. H., Cotrim da Cunha, L., Cox, P. M., Eliseev, A. V., Henson, S., Ishii, M., Jaccard, S., Koven, C., Lohila, A., Patra, P. K., Piao, S., Rogelj, J., Syampungani, S., Zaehle, S., and Zickfeld, K.: Global

Carbon and other Biogeochemical Cycles and Feedbacks, edited by: Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B.
505 R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, https://doi.org/10.1017/9781009157896.007, 2021.

Chen, D., Rojas, M., Samset, B. H., Cobb, K., Diongue Niang, A., Edwards, P., Emori, S., Faria, S. H., Hawkins, E., Hope, P., Huybrechts, P., Meinshausen, M., Mustafa, S. K., Plattner, G.-K., and Tréguier, A.-M.: Framing, Context, and Methods, edited
510 by: Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, https://doi.org/10.1017/9781009157896.003, 2021.

Cherchi, A., Fogli, P. G., Lovato, T., Peano, D., Iovino, D., Gualdi, S., Masina, S., Scoccimarro, E., Materia, S., Bellucci, A.,
515 and Navarra, A.: Global Mean Climate and Main Patterns of Variability in the CMCC-CM2 Coupled Model, Journal of Advances in Modeling Earth Systems, 11, 185–209, https://doi.org/10.1029/2018MS001369, 2019.

Church, J., Gregory, J., White, N., Platten, S., and Mitrovica, J.: Understanding and Projecting Sea Level Change, Oceanog., 24, 130–143, https://doi.org/10.5670/oceanog.2011.33, 2011.

Cocco, V., Joos, F., Steinacher, M., Frölicher, T. L., Bopp, L., Dunne, J., Gehlen, M., Heinze, C., Orr, J., Oschlies, A.,
520 Schneider, B., Segschneider, J., and Tjiputra, J.: Oxygen and indicators of stress for marine life in multi-model global warming projections, Biogeosciences, 10, 1849–1868, https://doi.org/10.5194/bg-10-1849-2013, 2013.

Danabasoglu, G., Lamarque, J. -F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S.,
525 Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kinnison, D., Kushner, P. J., Larson, V. E., Long, M. C., Mickelson, S., Moore, J. K., Nienhouse, E., Polvani, L., Rasch, P. J., and Strand, W. G.: The Community Earth System Model Version 2 (CESM2), J. Adv. Model. Earth Syst., 12, https://doi.org/10.1029/2019MS001916, 2020.

Deutsch, C., Ferrel, A., Seibel, B., Pörtner, H.-O., and Huey, R. B.: Climate change tightens a metabolic constraint on marine
530 habitats, Science, 348, 1132–1135, https://doi.org/10.1126/science.aaa1605, 2015.

Deutsch, C., Penn, J. L., and Seibel, B.: Metabolic trait diversity shapes marine biogeography, Nature, 585, 557–562, https://doi.org/10.1038/s41586-020-2721-y, 2020.

Diaz, R. J. and Rosenberg, R.: Spreading Dead Zones and Consequences for Marine Ecosystems, Science, 321, 926–929, https://doi.org/10.1126/science.1156401, 2008.

535 Doney, S. C., Fabry, V. J., Feely, R. A., and Kleypas, J. A.: Ocean Acidification: The Other CO2 Problem, Annu. Rev. Mar. Sci., 1, 169–192, https://doi.org/10.1146/annurev.marine.010908.163834, 2009.

Doney, S. C., Busch, D. S., Cooley, S. R., and Kroeker, K. J.: The Impacts of Ocean Acidification on Marine Ecosystems and Reliant Human Communities, Annu. Rev. Environ. Resour., 45, 83–112, https://doi.org/10.1146/annurev-environ-012320-083019, 2020.

540 Dutkiewicz, S., Scott, J. R., and Follows, M. J.: Winners and losers: Ecological and biogeochemical changes in a warming ocean, Global Biogeochem. Cycles, 27, 463–477, https://doi.org/10.1002/gbc.20042, 2013.

Edwards, N. R., Willmott, A. J., and Killworth, P. D.: On the Role of Topography and Wind Stress on the Stability of the Thermohaline Circulation, Journal of Physical Oceanography, 28, 756–778, https://doi.org/10.1175/1520-0485(1998)028<0756:OTROTA>2.0.CO;2, 1998.

545 Ekau, W., Auel, H., Pörtner, H.-O., and Gilbert, D.: Impacts of hypoxia on the structure and processes in pelagic communities (zooplankton, macro-invertebrates and fish), Biogeosciences, 7, 1669–1699, https://doi.org/10.5194/bg-7-1669-2010, 2010.

Ekstrom, J. A., Suatoni, L., Cooley, S. R., Pendleton, L. H., Waldbusser, G. G., Cinner, J. E., Ritter, J., Langdon, C., Van Hooidonk, R., Gledhill, D., Wellman, K., Beck, M. W., Brander, L. M., Rittschof, D., Doherty, C., Edwards, P. E. T., and Portela, R.: Vulnerability and adaptation of US shellfisheries to ocean acidification, Nature Clim Change, 5, 207–214, 550 https://doi.org/10.1038/nclimate2508, 2015.

Fabry, V., McClintock, J., Mathis, J., and Grebmeier, J.: Ocean Acidification at High Latitudes: The Bellwether, Oceanog., 22, 160–171, https://doi.org/10.5670/oceanog.2009.105, 2009.

Fanning, A. F. and Weaver, A. J.: An atmospheric energy-moisture balance model: Climatology, interpentadal climate change, and coupling to an ocean general circulation model, Journal of Geophysical Research: Atmospheres, 101, 15111–15128, 555 https://doi.org/10.1029/96JD01017, 1996.

Fox-Kemper, B., Hewitt, H. T., Xiao, C., Aðalgeirsdóttir, G., Drijfhout, S. S., Edwards, T. L., Golledge, N. R., Hemer, M., Kopp, R. E., Krinner, G., Mix, A., Notz, D., Nowicki, S., Nurhati, I. S., Ruiz, L., Sallée, J.-B., Slangen, A. B. A., and Yu, Y.: Ocean, Cryosphere and Sea Level Change, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Masson-Delmotte, V., 560 Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., Cambridge University Press, Cambridge, UK and New York, NY, USA, https://doi.org/10.1017/9781009157896.011, 2021.

Fröb, F., Bourgeois, T., Goris, N., Schwinger, J., and Heinze, C.: Simulated Abrupt Shifts in Aerobic Habitats of Marine Species in the Past, Present, and Future, Earth's Future, 12, e2023EF004141, https://doi.org/10.1029/2023EF004141, 2024.

565 Fu, W., Randerson, J. T., and Moore, J. K.: Climate change impacts on net primary production (NPP) and export production(EP) regulated by increasing stratification and phytoplankton communitystructure in the CMIP5 models, Biogeosciences, 13, 5151–5170, https://doi.org/10.5194/bg-13-5151-2016, 2016.

Gattuso, J.-P., Magnan, A., Billé, R., Cheung, W. W. L., Howes, E. L., Joos, F., Allemand, D., Bopp, L., Cooley, S. R., Eakin, C. M., Hoegh-Guldberg, O., Kelly, R. P., Pörtner, H.-O., Rogers, A. D., Baxter, J. M., Laffoley, D., Osborn, D., Rankovic, A., 570 Rochette, J., Sumaila, U. R., Treyer, S., and Turley, C.: Contrasting futures for ocean and society from different anthropogenic CO2 emissions scenarios, Science, 349, aac4722, https://doi.org/10.1126/science.aac4722, 2015.

Gimenez, I., Waldbusser, G. G., and Hales, B.: Ocean acidification stress index for shellfish (OASIS): Linking Pacific oyster larval survival and exposure to variable carbonate chemistry regimes, Elementa: Science of the Anthropocene, 6, 51, https://doi.org/10.1525/elementa.306, 2018.

575 Gruber, N.: Warming up, turning sour, losing breath: ocean biogeochemistry under global change, Phil. Trans. R. Soc. A., 369, 1980–1996, https://doi.org/10.1098/rsta.2011.0003, 2011.

Guinotte, J. M., Buddemeier, R. W., and Kleypas, J. A.: Future coral reef habitat marginality: temporal and spatial effects of climate change in the Pacific basin, Coral Reefs, 22, 551–558, https://doi.org/10.1007/s00338-003-0331-4, 2003.

Guinotte, J. M., Orr, J., Cairns, S., Freiwald, A., Morgan, L., and George, R.: Will human-induced changes in seawater
580 chemistry alter the distribution of deep-sea scleractinian corals?, Frontiers in Ecology and the Environment, 4, 141–146,
https://doi.org/10.1890/1540-9295(2006)004[0141:WHCISC]2.0.CO;2, 2006.

Hajima, T., Watanabe, M., Yamamoto, A., Tatebe, H., Noguchi, M. A., Abe, M., Ohgaito, R., Ito, A., Yamazaki, D., Okajima,
H., Ito, A., Takata, K., Ogochi, K., Watanabe, S., and Kawamiya, M.: Development of the MIROC-ES2L Earth system model
and the evaluation of biogeochemical processes and feedbacks, Geosci. Model Dev., 13, 2197–2244,
585 https://doi.org/10.5194/gmd-13-2197-2020, 2020.

Hameau, A., Frölicher, T. L., Mignot, J., and Joos, F.: Is deoxygenation detectable before warming in the thermocline?,
Biogeosciences, 17, 1877–1895, https://doi.org/10.5194/bg-17-1877-2020, 2020.

Heinze, C., Blenckner, T., Martins, H., Rusiecka, D., Döscher, R., Gehlen, M., Gruber, N., Holland, E., Hov, Ø., Joos, F.,
Matthews, J. B. R., Rødven, R., and Wilson, S.: The quiet crossing of ocean tipping points, Proc. Natl. Acad. Sci. U.S.A., 118,
590 e2008478118, https://doi.org/10.1073/pnas.2008478118, 2021.

Hermans, T. H. J., Gregory, J. M., Palmer, M. D., Ringer, M. A., Katsman, C. A., and Slangen, A. B. A.: Projecting Global
Mean Sea-Level Change Using CMIP6 Models, Geophysical Research Letters, 48, https://doi.org/10.1029/2020GL092064,
2021.

Hinkel, J., Lincke, D., Vafeidis, A. T., Perrette, M., Nicholls, R. J., Tol, R. S. J., Marzeion, B., Fettweis, X., Ionescu, C., and
595 Levermann, A.: Coastal flood damage and adaptation costs under 21st century sea-level rise, Proc. Natl. Acad. Sci. U.S.A.,
111, 3292–3297, https://doi.org/10.1073/pnas.1222469111, 2014.

Hoegh-Guldberg, O., Mumby, P. J., Hooten, A. J., Steneck, R. S., Greenfield, P., Gomez, E., Harvell, C. D., Sale, P. F.,
Edwards, A. J., Caldeira, K., Knowlton, N., Eakin, C. M., Iglesias-Prieto, R., Muthiga, N., Bradbury, R. H., Dubi, A., and
Hatziolos, M. E.: Coral Reefs Under Rapid Climate Change and Ocean Acidification, Science, 318, 1737–1742,
600 https://doi.org/10.1126/science.1152509, 2007.

Hunke, E. C. and Dukowicz, J. K.: An Elastic–Viscous–Plastic Model for Sea Ice Dynamics, J. Phys. Oceanogr., 27, 1849–
1867, https://doi.org/10.1175/1520-0485(1997)027<1849:AEVPMF>2.0.CO;2, 1997.

Jeltsch-Thömmes, A., Stocker, T. F., and Joos, F.: Hysteresis of the Earth system under positive and negative CO_2 emissions,
Environ. Res. Lett., 15, 124026, https://doi.org/10.1088/1748-9326/abc4af, 2020.

605 Jeltsch-Thömmes, A., Tran, G., Lienert, S., Keller, D. P., Oschlies, A., and Joos, F.: Earth system responses to carbon dioxide
removal as exemplified by ocean alkalinity enhancement: tradeoffs and lags, Environ. Res. Lett., 19, 054054,
https://doi.org/10.1088/1748-9326/ad4401, 2024.

Keller, D. P., Oschlies, A., and Eby, M.: A new marine ecosystem model for the University of Victoria Earth System Climate
Model, Geoscientific Model Development, 5, 1195–1220, https://doi.org/10.5194/gmd-5-1195-2012, 2012.

610 Kennedy, M. C. and O'Hagan, A.: Bayesian calibration of computer models, Journal of the Royal Statistical Society: Series B
(Statistical Methodology), 63, 425–464, https://doi.org/10.1111/1467-9868.00294, 2001.

Kleypas, J. A., Mcmanus, J. W., and Meñez, L. A. B.: Environmental Limits to Coral Reef Development: Where Do We Draw
the Line?, Am Zool, 39, 146–159, https://doi.org/10.1093/icb/39.1.146, 1999.

Kordas, R. L., Harley, C. D. G., and O'Connor, M. I.: Community ecology in a warming world: The influence of temperature on interspecific interactions in marine systems, J. Exp. Mar. Biol. Ecol., 400, 218–226, https://doi.org/10.1016/j.jembe.2011.02.029, 2011.

Kriegler, E., Hall, J. W., Held, H., Dawson, R., and Schellnhuber, H. J.: Imprecise probability assessment of tipping points in the climate system, Proc. Natl. Acad. Sci. U.S.A., 106, 5041–5046, https://doi.org/10.1073/pnas.0809117106, 2009.

Kwiatkowski, L., Torres, O., Bopp, L., Aumont, O., Chamberlain, M., Christian, J. R., Dunne, J. P., Gehlen, M., Ilyina, T., John, J. G., Lenton, A., Li, H., Lovenduski, N. S., Orr, J. C., Palmieri, J., Santana-Falcón, Y., Schwinger, J., Séférian, R., Stock, C. A., Tagliabue, A., Takano, Y., Tjiputra, J., Toyama, K., Tsujino, H., Watanabe, M., Yamamoto, A., Yool, A., and Ziehn, T.: Twenty-first century ocean warming, acidification, deoxygenation, and upper-ocean nutrient and primary production decline from CMIP6 model projections, Biogeosciences, 17, 3439–3470, https://doi.org/10.5194/bg-17-3439-2020, 2020.

Lauvset, S. K., Key, R. M., Olsen, A., van Heuven, S., Velo, A., Lin, X., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterström, S., Steinfeldt, R., Jeansson, E., Ishii, M., Perez, F. F., Suzuki, T., and Watelet, S.: A new global interior ocean mapped climatology: the 1◦ × 1◦ GLODAP version 2, 16, 2016.

Lenton, T. M.: Arctic Climate Tipping Points, AMBIO, 41, 10–22, https://doi.org/10.1007/s13280-011-0221-x, 2012.

Lenton, T. M., Held, H., Kriegler, E., Hall, J. W., Lucht, W., Rahmstorf, S., and Schellnhuber, H. J.: Tipping elements in the Earth's climate system, Proceedings of the National Academy of Sciences, 105, 1786–1793, https://doi.org/10.1073/pnas.0705414105, 2008.

Lienert, S. and Joos, F.: A Bayesian ensemble data assimilation to constrain model parameters and land-use carbon emissions, Biogeosciences, 15, 2909–2930, https://doi.org/10.5194/bg-15-2909-2018, 2018.

Limburg, K. E., Breitburg, D., Swaney, D. P., and Jacinto, G.: Ocean Deoxygenation: A Primer, One Earth, 2, 24–29, https://doi.org/10.1016/j.oneear.2020.01.001, 2020.

Locarnini, R. A., Mishonov, A. V., Antonov, J. I., Boyer, T. P., Garcia, H. E., Baranova, O. K., Zweng, M. M., Paver, C. R., Reagan, J. R., Johnson, D. R., Hamilton, M., Seidov, D., 1948-, and Levitus, S.: World ocean atlas 2013. Volume 1, Temperature, NOAA atlas NESDIS 73, https://doi.org/10.7289/V55X26VD, 2013.

MacDougall, A. H., Swart, N. C., and Knutti, R.: The Uncertainty in the Transient Climate Response to Cumulative CO2 Emissions Arising from the Uncertainty in Physical Climate Parameters, Journal of Climate, 30, 813–827, https://doi.org/10.1175/JCLI-D-16-0205.1, 2017.

Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B. (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, , https://doi.org/10.1017/9781009157896, 2021.

Mckay, M. D., Beckman, R. J., and Conover, W. J.: A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code, Technometrics, 42, 55–61, https://doi.org/10.1080/00401706.2000.10485979, 2000.

Meissner, K. J., Weaver, A. J., Matthews, H. D., and Cox, P. M.: The role of land surface dynamics in glacial inception: a study with the UVic Earth System Model, Climate Dynamics, 21, 515–537, https://doi.org/10.1007/s00382-003-0352-2, 2003.

650   Mengis, N., Keller, D. P., MacDougall, A. H., Eby, M., Wright, N., Meissner, K. J., Oschlies, A., Schmittner, A., MacIsaac, A. J., Matthews, H. D., and Zickfeld, K.: Evaluation of the University of Victoria Earth System Climate Model version 2.10 (UVic ESCM 2.10), Geoscientific Model Development, 13, 4183–4204, https://doi.org/10.5194/gmd-13-4183-2020, 2020.

Morée, A. L., Clarke, T. M., Cheung, W. W. L., and Frölicher, T. L.: Impact of deoxygenation and warming on global marine species in the 21st century, Biogeosciences, 20, 2425–2454, https://doi.org/10.5194/bg-20-2425-2023, 2023.

655   Müller, S. A., Joos, F., Edwards, N. R., and Stocker, T. F.: Water Mass Distribution and Ventilation Time Scales in a Cost-Efficient, Three-Dimensional Ocean Model, Journal of Climate, 19, 5479–5499, https://doi.org/10.1175/JCLI3911.1, 2006.

Nijsse, F. J. M. M., Cox, P. M., and Williamson, M. S.: Emergent constraints on transient climate response (TCR) and equilibrium climate sensitivity (ECS) from historical warming in CMIP5 and CMIP6 models, Earth Syst. Dynam., 11, 737–750, https://doi.org/10.5194/esd-11-737-2020, 2020.

660   Oliver, E. C. J., Benthuysen, J. A., Darmaraki, S., Donat, M. G., Hobday, A. J., Holbrook, N. J., Schlegel, R. W., and Sen Gupta, A.: Marine Heatwaves, Annu. Rev. Mar. Sci., 13, 313–342, https://doi.org/10.1146/annurev-marine-032720-095144, 2021.

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., and Sanderson, B. M.: The Scenario Model Intercomparison Project 665   (ScenarioMIP) for CMIP6, Geosci. Model Dev., 9, 3461–3482, https://doi.org/10.5194/gmd-9-3461-2016, 2016.

O'Neill, B. C., Oppenheimer, M., Warren, R., Hallegatte, S., Kopp, R. E., Pörtner, H. O., Scholes, R., Birkmann, J., Foden, W., Licker, R., Mach, K. J., Marbaix, P., Mastrandrea, M. D., Price, J., Takahashi, K., Van Ypersele, J.-P., and Yohe, G.: IPCC reasons for concern regarding climate change risks, Nature Clim Change, 7, 28–37, https://doi.org/10.1038/nclimate3179, 2017.

670   Orr, J. C. and Epitalon, J.-M.: Improved routines to model the ocean carbonate system: mocsy 2.0, Geosci. Model Dev., 8, 485–499, https://doi.org/10.5194/gmd-8-485-2015, 2015.

Orr, J. C., Fabry, V. J., Aumont, O., Bopp, L., Doney, S. C., Feely, R. A., Gnanadesikan, A., Gruber, N., Ishida, A., Joos, F., Key, R. M., Lindsay, K., Maier-Reimer, E., Matear, R., Monfray, P., Mouchet, A., Najjar, R. G., Plattner, G.-K., Rodgers, K. B., Sabine, C. L., Sarmiento, J. L., Schlitzer, R., Slater, R. D., Totterdell, I. J., Weirig, M.-F., Yamanaka, Y., and Yool, A.: 675   Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms, Nature, 437, 681–686, https://doi.org/10.1038/nature04095, 2005.

Pacanowski, R. C. E.: MOM 2. Documentation, User's Guide and Reference Manual. Technical Report 3.2. GFDL Ocean Group, GFDL, Princeton, New Jersey, 1996.

Peng, G., Matthews, J. L., Wang, M., Vose, R., and Sun, L.: What Do Global Climate Models Tell Us about Future Arctic Sea 680   Ice Coverage Changes?, Climate, 8, 15, https://doi.org/10.3390/cli8010015, 2020.

Pinsky, M. L., Selden, R. L., and Kitchel, Z. J.: Climate-Driven Shifts in Marine Species Ranges: Scaling from Organisms to Communities, Annu. Rev. Mar. Sci., 12, 153–179, https://doi.org/10.1146/annurev-marine-010419-010916, 2020.

Pörtner, H.-O., Roberts, D. C., Masson-Delmotte, V., Zhai, P., Tignor, M. M. B., Poloczanska, E., Mintenbeck, K., Alegría, A., Nicolai, M., Okem, A. E., Petzold, J., Rama, B., and Weyer, N. M. (Eds.): IPCC Special Report on the Ocean and 685   Cryosphere in a Changing Climate, Cambridge University Press, Cambridge, UK and New York, NY, USA, 755 pp., 2019.

Rasmussen, C. E. and Williams, C. K. I.: Gaussian Processes for Machine Learning, The MIT Press, https://doi.org/10.7551/mitpress/3206.001.0001, 2005.

Ritz, S. P., Stocker, T. F., and Joos, F.: A Coupled Dynamical Ocean–Energy Balance Atmosphere Model for Paleoclimate Studies, Journal of Climate, 24, 349–375, 2011.

690   Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P.: Design and Analysis of Computer Experiments, Statistical Science, 4, 409–423, 1989.

Samanta, A., Anderson, B. T., Ganguly, S., Knyazikhin, Y., Nemani, R. R., and Myneni, R. B.: Physical Climate Response to a Reduction of Anthropogenic Climate Forcing, Earth Interactions, 14, 1–11, https://doi.org/10.1175/2010EI325.1, 2010.

Santana-Falcón, Y., Yamamoto, A., Lenton, A., Jones, C. D., Burger, F. A., John, J. G., Tjiputra, J., Schwinger, J., Kawamiya,
695   M., Frölicher, T. L., Ziehn, T., and Séférian, R.: Irreversible loss in marine ecosystem habitability after a temperature overshoot, Commun Earth Environ, 4, 343, https://doi.org/10.1038/s43247-023-01002-1, 2023.

Schwinger, J., Asaadi, A., Goris, N., and Lee, H.: Possibility for strong northern hemisphere high-latitude cooling under negative emissions, Nat Commun, 13, 1095, https://doi.org/10.1038/s41467-022-28573-5, 2022.

Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., Colin, J., Decharme, B., Delire, C., Berthet, S., Chevallier,
700   M., Sénési, S., Franchisteguy, L., Vial, J., Mallet, M., Joetzjer, E., Geoffroy, O., Guérémy, J., Moine, M., Msadek, R., Ribes, A., Rocher, M., Roehrig, R., Salas-y-Mélia, D., Sanchez, E., Terray, L., Valcke, S., Waldman, R., Aumont, O., Bopp, L., Deshayes, J., Éthé, C., and Madec, G.: Evaluation of CNRM Earth System Model, CNRM-ESM2-1: Role of Earth System Processes in Present-Day and Future Climate, J. Adv. Model. Earth Syst., 11, 4182–4227, https://doi.org/10.1029/2019MS001791, 2019.

705   Seland, Ø., Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y.-C., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Karset, I. H. H., Landgren, O., Liakka, J., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, T., and Schulz, M.: Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations, Geosci. Model Dev., 13, 6165–6200, https://doi.org/10.5194/gmd-13-6165-2020, 2020.

710   Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., O'Connor, F. M., Stringer, M., Hill, R., Palmieri, J., Woodward, S., Mora, L., Kuhlbrodt, T., Rumbold, S. T., Kelley, D. I., Ellis, R., Johnson, C. E., Walton, J., Abraham, N. L., Andrews, M. B., Andrews, T., Archibald, A. T., Berthou, S., Burke, E., Blockley, E., Carslaw, K., Dalvi, M., Edwards, J., Folberth, G. A., Gedney, N., Griffiths, P. T., Harper, A. B., Hendry, M. A., Hewitt, A. J., Johnson, B., Jones, A., Jones, C. D., Keeble, J., Liddicoat, S., Morgenstern, O., Parker, R. J., Predoi, V., Robertson, E., Siahaan, A., Smith, R. S., Swaminathan,
715   R., Woodhouse, M. T., Zeng, G., and Zerroukat, M.: UKESM1: Description and Evaluation of the U.K. Earth System Model, J. Adv. Model. Earth Syst., 11, 4513–4558, https://doi.org/10.1029/2019MS001739, 2019.

Steinacher, M. and Joos, F.: Transient Earth system responses to cumulative carbon dioxide emissions: linearities, uncertainties, and probabilities in an observation-constrained model ensemble, Biogeosciences, 13, 1071–1103, https://doi.org/10.5194/bg-13-1071-2016, 2016.

720   Steinacher, M., Joos, F., Frölicher, T. L., Plattner, G.-K., and Doney, S. C.: Imminent ocean acidification in the Arctic projected with the NCAR global coupled carbon cycle-climate model, Biogeosciences, 6, 515–533, https://doi.org/10.5194/bg-6-515-2009, 2009.

Steinacher, M., Joos, F., and Stocker, T. F.: Allowable carbon emissions lowered by multiple climate targets, Nature, 499, 197–201, https://doi.org/10.1038/nature12269, 2013.

725 Stroeve, J. C., Kattsov, V., Barrett, A., Serreze, M., Pavlova, T., Holland, M., and Meier, W. N.: Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations, Geophys. Res. Lett., 39, 2012GL052676, https://doi.org/10.1029/2012GL052676, 2012.

Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Hanna, S., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., von
730 Salzen, K., Yang, D., and Winter, B.: The Canadian Earth System Model version 5 (CanESM5.0.3), Geosci. Model Dev., 12, 4823–4873, https://doi.org/10.5194/gmd-12-4823-2019, 2019.

Terhaar, J., Kwiatkowski, L., and Bopp, L.: Emergent constraint on Arctic Ocean acidification in the twenty-first century, Nature, 582, 379–383, https://doi.org/10.1038/s41586-020-2360-3, 2020.

Terhaar, J., Torres, O., Bourgeois, T., and Kwiatkowski, L.: Arctic Ocean acidification over the 21st century co-driven by
735 anthropogenic carbon increases and freshening in the CMIP6 model ensemble, Biogeosciences, 18, 2221–2240, https://doi.org/10.5194/bg-18-2221-2021, 2021.

Tittensor, D. P., Novaglio, C., Harrison, C. S., Heneghan, R. F., Barrier, N., Bianchi, D., Bopp, L., Bryndum-Buchholz, A., Britten, G. L., Büchner, M., Cheung, W. W. L., Christensen, V., Coll, M., Dunne, J. P., Eddy, T. D., Everett, J. D., Fernandes-Salvador, J. A., Fulton, E. A., Galbraith, E. D., Gascuel, D., Guiet, J., John, J. G., Link, J. S., Lotze, H. K., Maury, O., Ortega-
740 Cisneros, K., Palacios-Abrantes, J., Petrik, C. M., du Pontavice, H., Rault, J., Richardson, A. J., Shannon, L., Shin, Y.-J., Steenbeek, J., Stock, C. A., and Blanchard, J. L.: Next-generation ensemble projections reveal higher climate risks for marine ecosystems, Nat. Clim. Chang., 11, 973–981, https://doi.org/10.1038/s41558-021-01173-9, 2021.

Tokarska, K. B., Stolpe, M. B., Sippel, S., Fischer, E. M., Smith, C. J., Lehner, F., and Knutti, R.: Past warming trend constrains future warming in CMIP6 models, Sci. Adv., 6, eaaz9549, https://doi.org/10.1126/sciadv.aaz9549, 2020.

745 UNFCCC: Adoption of the Paris Agreement, Report No. FCCC/CP/2015/L.9/Rev.1, 2015.

Van Gennip, S. J., Popova, E. E., Yool, A., Pecl, G. T., Hobday, A. J., and Sorte, C. J. B.: Going with the flow: the role of ocean circulation in global marine ecosystems under a changing climate, Glob. Chang. Biol., 23, 2602–2617, https://doi.org/10.1111/gcb.13586, 2017.

Van Westen, R. M., Kliphuis, M., and Dijkstra, H. A.: Physics-based early warning signal shows that AMOC is on tipping
750 course, Sci. Adv., 10, eadk1189, https://doi.org/10.1126/sciadv.adk1189, 2024.

Vaquer-Sunyer, R. and Duarte, C. M.: Thresholds of hypoxia for marine biodiversity, Proc. Natl. Acad. Sci. U.S.A., 105, 15452–15457, https://doi.org/10.1073/pnas.0803833105, 2008.

Waldbusser, G. G., Hales, B., Langdon, C. J., Haley, B. A., Schrader, P., Brunner, E. L., Gray, M. W., Miller, C. A., and Gimenez, I.: Saturation-state sensitivity of marine bivalve larvae to ocean acidification, Nature Clim Change, 5, 273–280,
755 https://doi.org/10.1038/nclimate2479, 2015.

WCRP Global Sea Level Budget Group: Global sea-level budget 1993–present, Earth Syst. Sci. Data, 10, 1551–1590, https://doi.org/10.5194/essd-10-1551-2018, 2018.

Weaver, A. J., Eby, M., Wiebe, E. C., Bitz, C. M., Duffy, P. B., Ewen, T. L., Fanning, A. F., Holland, M. M., MacFadyen, A., Matthews, H. D., Meissner, K. J., Saenko, O., Schmittner, A., Wang, H., and Yoshimori, M.: The UVic earth system climate
760 model: Model description, climatology, and applications to past, present and future climates, Atmosphere-Ocean, 39, 361–428, https://doi.org/10.1080/07055900.2001.9649686, 2001.

Weaver, A. J., Sedláček, J., Eby, M., Alexander, K., Crespin, E., Fichefet, T., Philippon-Berthier, G., Joos, F., Kawamiya, M., Matsumoto, K., Steinacher, M., Tachiiri, K., Tokos, K., Yoshimori, M., and Zickfeld, K.: Stability of the Atlantic meridional overturning circulation: A model intercomparison, Geophys. Res. Lett., 39, 2012GL053763, https://doi.org/10.1029/2012GL053763, 2012.

Weijer, W., Cheng, W., Garuba, O. A., Hu, A., and Nadiga, B. T.: CMIP6 Models Predict Significant 21st Century Decline of the Atlantic Meridional Overturning Circulation, Geophys. Res. Lett., 47, e2019GL086075, https://doi.org/10.1029/2019GL086075, 2020.

Williamson, D., Blaker, A. T., Hampton, C., and Salter, J.: Identifying and removing structural biases in climate models with history matching, Clim Dyn, 45, 1299–1324, https://doi.org/10.1007/s00382-014-2378-z, 2015.

Zickfeld, K., Levermann, A., Morgan, M. G., Kuhlbrodt, T., Rahmstorf, S., and Keith, D. W.: Expert judgements on the response of the Atlantic meridional overturning circulation to climate change, Climatic Change, 82, 235–265, https://doi.org/10.1007/s10584-007-9246-3, 2007.

Ziehn, T., Chamberlain, M. A., Law, R. M., Lenton, A., Bodman, R. W., Dix, M., Stevens, L., Wang, Y.-P., and Srbinovsky, J.: The Australian Earth System Model: ACCESS-ESM1.5, JSHESS, 70, 193, https://doi.org/10.1071/ES19035, 2020.

Zweng, M. M., Reagan, J. R., Antonov, J. I., Locarnini, R. A., Mishonov, A. V., Boyer, T. P., Garcia, H. E., Baranova, O. K., Johnson, D. R., Seidov, D., 1948-, Biddle, M. M., and Levitus, S.: World ocean atlas 2013. Volume 2, Salinity, NOAA Atlas NESDIS 74, https://doi.org/10.7289/V5251G4D, 2013.