CMIP6 simulations with the compact Earth system model OSCAR v3.1

Yann Quilcaille[1*], Thomas Gasser[2], Philippe Ciais[3], Olivier Boucher[4]

[1] International Institute for Applied Systems Analysis (IIASA), 2361 Laxenburg, Austria
* now at Institute for Atmospheric and Climate Science, Department of Environmental Systems Sciences, ETH Zürich, Zürich, Switzerland
[2] International Institute for Applied Systems Analysis (IIASA), 2361 Laxenburg, Austria
[3] Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, Université Paris-Saclay, CEA – CNRS – UVSQ, 91191 Gif-sur-Yvette, France
[4] Institut Pierre-Simon Laplace, Sorbonne Université, CNRS 75252 Paris, France

Correspondence to: Yann Quilcaille (yann.quilcaille@env.ethz.ch)

Abstract. Reduced complexity models, also called simple climate models or compact models, provide an alternative to Earth system models (ESMs) with lower computational costs, although at the expense of spatial and temporal information. It remains important to evaluate and validate these reduced complexity models. Here, we evaluate ~~the newest~~a recent version (v3.1) of the OSCAR model using observations and results from ESMs from the current Coupled Model Intercomparison Project 6 (CMIP6). The results follow the same post-processing used for the contribution of OSCAR to the Reduced Complexity Model Intercomparison Project Phase 2~~, with~~ regarding the identification of stable configurations and the use of observational constraints. ~~Overall, OSCAR v3.1 shows good agreement with observations, ESMs and emerging properties.~~These constraints succeed in decreasing the ~~increase~~overestimation of~~in~~ global surface air temperature over 2000-2019 with reference to 1961-1900 from 0.60 ± 0.11 K to 0.55 ± 0.04 K (~~,~~ the ~~target~~constraint being 0.54 ± 0.05 K). The Equilibrium Climate Sensitivity (ECS) of the unconstrained OSCAR is 3.17 ± 0.63 K, while CMIP5 & CMIP6 models have ~~respectively~~ ECSs of 3.2 ± 0.7 K and 3.7 ± 1.1 K, respectively. Applying observational constraints to OSCAR reduces the ECS to 2.78 ± 0.47K. Overall, ~~It~~the model qualitatively reproduces the responses of complex ESMs, ~~for all aspects of the Earth system. We observe~~although ~~S~~some ~~of the~~ differences ~~with these models~~remain~~, most of them being~~ are due to the impact of observational constraints on the weighting of parametrizations. Specific features of OSCAR also contribute to these differences, such as its fully interactive atmospheric chemistry and endogenous calculations of biomass burning, wetlands CH_4 and permafrost CH_4 and CO_2 emissions. Identified main points of needed improvements of the OSCAR model include a low sensitivity of the land carbon cycle to climate change, an instability of the ocean carbon cycle, the seemingly too simple climate module, and the too strong climate feedback involving short-lived species. Beyond providing a key diagnosis of the OSCAR model in the context of the reduced-complexity models ~~intercomparison project (RCMIP)~~, this work is also meant to help with the upcoming calibration of OSCAR on CMIP6 results, and to provide a large group of CMIP6 simulations run consistently within a probabilistic framework.

1. Introduction

Complex models such as Earth system models (ESMs) are used for climate projections (Collins et al., 2013). ESMs provide gridded detailed process-based outputs (Flato et al., 2013), but these strengths are mitigated by heavy computational costs. As a complement, reduced-complexity models, also called simple climate models (SCMs), prove useful to investigate couplings and uncertainties (Nicholls et al., 2020; Clarke et al., 2014), especially for large ensembles of scenarios and statistical analysis of uncertainties to model parameters (Gasser et al., 2015; Li et al., 2016; Quilcaille et al., 2018). SCMs run significantly faster, thanks to a parametric modelling approach often calibrated on more complex models such as ESMs (Meinshausen et al., 2011a; Crichton et al., 2014; Hartin et al., 2015; Gasser et al., 2017; Smith et al., 2018; Dorheim et al., 2021). Although simpler than ESMs, those models exhibit diversity in their modelling and calibration (Nicholls et al., 2020; Nicholls et al., 2021). Reduced complexity models need to be validated despite their calibration and their relative simplicity. Reduced complexity models are often built as a combination of modules, each dedicated to aspects of the Earth system, such as the atmospheric chemistry, the oceanic carbon cycle, the climate response to radiative forcings, etc. These models may be developed as unique emulators, with all modules calibrated together to emulate a single ESM. They may also be developed as a combination of emulators, with each module calibrated separately, as it is the case for OSCAR. ~~Under such an approach, the range of potential modelling outcomes is broadened but the need for validation is increased.~~

Thanks to its relative simplicity, OSCAR is capable to easily include additional processes using existing models of higher complexity (Gasser et al., 2018). This SCM is designed to run in a probabilistic framework, where every ensemble member corresponds to the parametrization of these processes. Thus, OSCAR combines features from a large set of models (Gasser et al., 2017): for instance, emissions from land-use change, permafrost, wetlands and biomass burning are endogenously calculated in the model. Under such an approach, the range of potential modelling outcomes is broader than that of the ESMs. Yet, it also increases the need for validation. As a potential correction, OSCAR may also easily integrate observational constraints. In this paper, we evaluate this modelling chain.

~~In this paper, experiments~~ Experiments designed under the Coupled Model Intercomparison Project 6 (Eyring et al., 2016) are used to ~~diagnose~~ evaluate the performance~~s~~ of ~~the latest~~ version 3.1 of OSCAR, comparing its results to observations and other model outputs. We briefly describe the model and its update, the probabilistic setup used, and how it has been constrained using observations. We present the CMIP6 simulations run with OSCAR and compare their results to the available CMIP6 ESM runs. Beyond ~~diagnosis~~ evaluation and despite being a simple model, OSCAR has a number of specificities that make it interesting to some ~~of~~ CMIP6-endorsed MIPs: CDRMIP (Keller et al., 2018a) and ZECMIP (Jones et al., 2019) thanks to its advanced carbon cycle, and LUMIP (Lawrence et al., 2016) thanks to its book-keeping land use module. OSCAR is also part of the RCMIP project phases 1 and 2 (Nicholls et al., 2020; Nicholls et al., 2021), whose objective is to compare reduced complexity models together and against CMIP6 and CMIP5 simulations.

In this study, we focus on several aspects of the model. To begin with, we describe OSCAR already detailed in (Gasser et al., 2017) and the setup~~the approach based on the exclusion of diverging parametrizations and observational constraints is~~

~~only briefly analysed, for it is the one~~ that was used in RCMIP phase 2 (Nicholls et al., 2021). The climate response of the model is investigated using ~~I~~idealized experiments from the DECK and RCMIP (Nicholls et al., 2020)~~ are used to evaluate the climate response, while o~~. The carbon response is then analyzed as well thanks to other idealized experiments from the DECK and C4MIP (Jones et al., 2016)~~ to evaluate the carbon cycle response~~. The performances of OSCAR to reconstruct the historical period are evaluated using~~Then, we use~~ experiments from the DECK (Eyring et al., 2016)~~ to simulate climate change over the historical period~~. We extend this analysis thanks to an attribution exercise of historical global temperature change, ~~Experiments~~ based on experiments from DAMIP (Gillett et al., 2016), AerChemMIP (Collins et al., 2017), C4MIP (Jones et al., 2016) and LUMIP~~ (Lawrence et al., 2016) form the basis for an attribution exercise of historical global temperature change~~. Comparison on climate projections are then obtained using ScenarioMIP (O'Neill et al., 2016). Insights are calculated on the zero emission committed warming using ZECMIP (Jones et al., 2019). Further analysis on the behaviour of OSCAR is provided in the Appendix B.

2. Experimental setup

2.1. Brief description of OSCAR v3.1

OSCAR v3.1 is an open-source Earth system model of reduced complexity, whose modules mimic models of higher complexity~~, and~~. OSCAR is meant to be used in a probabilistic fashion (Gasser et al., 2017). A conceptual description of OSCAR v3.1 is given in Figure 1. The full description of OSCAR v2.2 can be found in (Gasser et al., 2017), providing details on its structure, equations and calibration. Changes from v2.2 to v3.1 are detailed in (Gasser et al., 2020a).

Global surface temperature changes in response to radiative forcing follows a two-box model formulation (Geoffroy et al., 2013b). Global precipitation is deduced from global surface temperature and the atmospheric fraction of radiative forcing (Shine et al., 2015). Linear scaling on the global variables is used to estimate regional temperature and precipitation changes, over five broad world regions (IIASA, 2018a). OSCAR calculates the radiative forcing caused by greenhouse gases (CO_2, CH_4, N_2O, 37 halogenated compounds), short-lived climate forcers (tropospheric and stratospheric ozone, stratospheric water vapour, nitrates, sulphates, black carbon, primary and secondary organic aerosols) and changes in surface albedo.

The ocean carbon cycle is based on the mixed-layer response function of (Joos et al., 1996), albeit with an added stratification of the upper ocean derived from CMIP5 (Arora et al., 2013) and with an updated carbonate chemistry. The land carbon cycle is divided into five biomes and the same five regions as previously~~, and~~. E~~-e~~ach of the 25 biome/region combinations follows a three-box model (soil, litter and vegetation) described by (Gasser et al., 2020a). The preindustrial state of the land carbon cycle is calibrated against TRENDYv7 (Le Quéré et al., 2018a) and its transient response to CO_2 and climate is calibrated against CMIP5 models (Arora et al., 2013).

~~Additionally,~~ OSCAR endogenously estimates key aspects of the carbon cycle. A dedicated book-keeping module tracks land cover change, wood harvest and shifting cultivation, which allows OSCAR to estimate its own CO_2 emissions from land-use change (Gasser et al., 2020a; Gasser and Ciais, 2013b). Permafrost thaw and the resulting emissions of CO_2 and

CH_4 are also accounted for (Gasser et al., 2018). CH_4 emissions from wetlands are calibrated on WETCHIMP (Melton et al., 2013). In addition, biomass burning emissions are calculated endogenously on the basis of the book-keeping module and wildfires that are simulated as part of the land carbon cycle (Gasser et al., 2017). The latter emissions were subtracted from the input data used to drive OSCAR to avoid double counting.

The atmospheric lifetimes of non-CO_2 greenhouse gases are impacted by non-linear tropospheric (Holmes et al., 2013) and stratospheric (Prather et al., 2015) chemistries. Tropospheric ozone follows the formulation by (Ehhalt et al., 2001) but recalibrated on ACCMIP (Stevenson et al., 2013). Stratospheric ozone is derived from (Newman et al., 2007) and (Ravishankara et al., 2009). Aerosol-radiation interactions are based on CMIP5 and AeroCom2 (Myhre et al., 2013), while aerosol-cloud interactions depend on the hydrophilic fraction of each aerosol and follows a logarithmic formulation (Hansen et al., 2005; Stevens, 2015). Surface albedo change induced by land-cover change follows (Bright and Kvalevåg, 2013). The impact of black carbon deposition on snow albedo is calibrated on ACCMIP globally (Lee et al., 2013) and regionalized following (Reddy and Boucher, 2007).

We pinpoint that OSCAR v3.1 is still calibrated on CMIP5 ESMs, and therefore not meant to emulate CMIP6 models. Furthermore, each module is calibrated on available models, but not all ESMs have implemented every aspect modelled in OSCAR, such as permafrost or biomass burning. It means that OSCAR does not emulate any given ESM, but it combines modules emulating specific parts of these models. Every parametrization of OSCAR is thus a combination of parameters, and some of these combinations may be unrealistic and need post-processing to keep only the physically realistic ones, as explained in section 2.3.

2.2. CMIP6 and RCMIP experiments

A total of 99 experiments were run with OSCAR, 75 being from CMIP6 and 24 from RCMIP. A list of these experiments is provided in Table 1~~Table 1~~. We selected~~ed~~ the experiments according to several criteria: typically, experiments are global and/or with long time-series of output requested, and experiments do not overly focus on a given process or short time scales. In addition, RCMIP requested~~ed~~ additional experiments to complement those of CMIP6, mostly extended and additional scenarios, including the RCP scenarios from the previous CMIP5 exercise (Meinshausen et al., 2011b). Between the CMIP5 and CMIP6 historical simulations, the concentration- and emission-driven ones, and the land-only experiments of LUMIP, eight different spin-up and control experiments had to be performed. Every spin-up is a recycling of the preindustrial forcing over 1000 years.

We use driving datasets for historical concentrations of greenhouse gases (Meinshausen et al., 2017), projected concentrations of greenhouse gases (ESGF, 2018), emissions (IIASA, 2018b; Gidden et al., 2019; Hoesly et al., 2018), land-use (LUH2, 2018), solar activity (Matthes et al., 2017), volcanic activity (Zanchettin et al., 2016) and the land-only climate climatology for LUMIP experiments (Lawrence et al., 2016). The extensions of scenarios are not those that were initially foreseen (O'Neill et al., 2016), but those that have effectively been used during the CMIP6 exercise (Meinshausen et al.,

2019). The volcanic aerosol optical depth has been treated to scale and extend AR5 volcanic radiative forcing (IPCC, 2013), to comply with the requirement of OSCAR to have a radiative forcing as driver for this contribution.

Every single experiment is run for 10,000 different configurations of OSCAR, drawn randomly from the pool of all possible parameters values in a Monte-Carlo setup (Gasser et al., 2017). Altogether, the combined experiments and Monte Carlo members sum to 569,700,000 simulated years.

2.3. Post-processing: exclusion and constraining

As described in (Gasser et al., 2017), most of the equations of OSCAR may use different sets of parameters or even different forms of equations. These parameters arise from the training over different models, while the forms of equations find their justification in the literature. Each combination of parameters and equations is defined as a configuration of OSCAR and represent a different possible model of the Earth system. A Monte Carlo setup is used with OSCAR over these configurations. This method for the uncertainty in the modelling of the Earth system comes with two side-effects: some combinations may be physically unrealistic, and some parameterizations may become numerically unstable when the model is pushed to the edge of the validity domain of its parametrizations. Therefore, the raw outputs of the simulations undergo two rounds of post-processing: one to exclude the diverging simulations, and one to constrain the resulting Monte Carlo ensemble. We ~~highlight~~ remind that the same exclusions and constraints are used for the contribution of OSCAR in RCMIP phase 2 (Nicholls et al., 2021). All details about the method are provided in Appendix A. All final outputs and results are provided as the resulting weighted means and standard deviations, using the normalised likelihood as weight. The effect of this constraining is further discussed in the next section.

~~In the exclusion round, we identify and discard the configurations that lead to a numerical divergence of the model as illustrated with Figure **2**. Every experiment undergoes a thorough search, and we developed heuristic criteria to exclude these diverging runs by trial-error. We identify divergences occurring in high warming scenarios, mostly when the oceanic carbon sink drops and then oscillates. We explain this instability with the stratification of the ocean surface, as detailed in equation 4 of (Gasser et al., 2017). Some parametrizations under high warming scenarios exhibit an additional mode, not diverging in the strictest sense, yet, with the ocean carbon sink becoming a source and then switching back to a sink, which we identified as a physically unrealistic behaviour of the parametrization.~~

~~To discard the unrealistic configurations, we use the experiments *ssp585*, *ssp370*, *1pctCO2* and *abrupt-4xCO2* for their high warming over different timescales. We use the ocean sink, the land sink, the CO_2 emissions from LUC and the CO_2 emissions from permafrost, to ensure that the whole carbon cycle remains within reasonable boundaries. The criteria are set based on the performances of the remaining subset. In general, we use 20 PgC/yr in absolute value as a threshold for divergence. Over *ssp585* and *ssp370*, the domain is restrained to strictly positive values, due to the additional mode mentioned previously. Over *abrupt-4xCO2*, the criteria are applied over the last 50 years of the experiments only. In~~

Field Code Changed

1pctCO2, the run is extended by another 100 years for better identification. As illustrated in Figure 2, most of the exclusions are related to ocean carbon sink, the other variables bring only little exclusions. The 1118 configurations not causing any divergences in all the experiments are kept as a common set of configurations for all experiments.

The need for exclusion is stronger as the atmospheric concentration of CO_2 and the global surface temperature increase. We acknowledge that when a significant fraction of the configurations is excluded, confidence in our model's result is lowered, but such a limitation of the validity domain is inherent to reduced-complexity models. The model's results might as well depend on the set thresholds for exclusions. However, this bias is reduced through the constraining round because configurations with unrealistic carbon cycles receive a low likelihood.

We observed that in most cases, the reason of the exclusion is due to a diverging ocean sink. The ocean carbon cycle of OSCAR is its oldest module (Gasser et al., 2017), and should be redesigned for more stable behaviour under high warming scenarios. A possibility is to increase the number of sub-timesteps in the oceanic carbon module to avoid this issue for a fraction of the configurations, but it comes at the expense of the computational cost of the model.

After this exclusion, the outputs of OSCAR are constrained using observations. As done for RCMIP phase 2 (Nicholls et al., 2021), the objective of this constraining round is to use the flexibility and the probabilistic frameworks of the reduced complexity models to synthesize lines of evidence with the modelling of the Earth system. With OSCAR, we assess the physical likelihood of the model's configurations using lines of evidence from the literature. For every constraint, we extend a method already used with OSCAR but with only one constraint (Gasser et al., 2020a; Le Quéré et al., 2018b). We assume a distribution from which we derive the likelihood of every configuration, as illustrated in equation A1 of (Gasser et al., 2020a). The product of the probabilities over the set of constraints is the final likelihood of the configurations.

As the first observational constraint, we choose the surface air-ocean blended temperature change over 2000-2019 with reference to 1961-1990 are used, provided as an assessed range by RCMIP (Nicholls et al., 2021) from the HadCRUT 4.6.0.0 dataset (Morice et al., 2012). This constraint is meant to provide information on the climate system. To constrain the carbon cycle, we use compatible fossil fuel emissions. For now, OSCAR v3.1 is calibrated on CMIP5, which motivates the use of the compatible emissions of CMIP5, not those of CMIP6. An initial set of constraints based solely on observations had revealed that using projections helped the overall constraining round, thanks to the larger perturbation in the scenarios than in the historical period. Thus we choose the CMIP5 cumulative compatible fossil fuel emissions over the concentrations-driven historical and 4 RCPs are used (Ciais et al., 2013b). To further constrain the partitioning of the carbon sinks between land and ocean, we use data on the cumulative net ocean to atmosphere flux of CO_2 over 1750-2011 (Ciais et al., 2013b).

All final outputs and results are provided as the resulting weighted means and standard deviations, using the normalised likelihood as weight. The effect of this constraining is further discussed in the next section.

Field Code Changed

Field Code Changed

3. ~~Diagnosis~~ Evaluation of OSCAR v3.1

In the previous~~former~~ section, we give an overview of OSCAR, explain which experiments are run and shortly describe how the results are processed. Given this experimental setup, we evaluate how OSCAR reproduces key features by comparing against other models and observations. We investigate the extent of the corrections brought by the constraints in Section 3.1. As the two main components of the Earth system, the climate and carbon cycle responses are then respectively investigated in Sections 3.2 and 3.3. We evaluate the capacity of OSCAR to reconstruct the historical period in Section 3.4 and calculate the contributions of individual forcings over the historical warming in Section 3.5. After evaluating the historical period, we evaluate how OSCAR performs on scenarios, comparing against ESMs in Section 3.6. The zero emissions commitment is ~~use~~presented in Section 3.7 ~~as a more advanced property of the Earth system~~ to compare the performance of OSCAR with respect to other models. Additional experiments are used to provide insights on the behaviour of OSCAR, albeit not used for evaluation of the model, as detailed in Section 3.8 and Appendix B.

~~3.~~

Formatted: Indent: First line: 0.63 cm

Formatted: Normal, No bullets or numbering

3.1. Effect of the constraints

Our constraining approach ~~markedly~~ corrects natural biases in OSCAR, as illustrated in ~~Figure 3~~Figure . The change in global surface air temperature (GSAT) over 2000-2019 with regard to 1961-1990 is constrained to a value of 0.54 ± 0.05 K. Without the constraint, OSCAR v3.1 reaches 0.60 ± 0.11 K. Due to the combination of observational constraints, OSCAR v3.1 is corrected to 0.55 ± 0.04 K.

Regarding the carbon cycle, the unconstrained OSCAR shows a negative bias in the cumulative net land carbon sink (i.e. a too weak removal), balanced by lower cumulative compatible fossil-fuel emissions. ~~Using o~~Observational constraints reduce~~s~~ these biases but do~~es~~ not entirely remove them. After applying the constraints, the uncertainty ranges of the net land flux and of fossil-fuel emissions are reduced. ~~The~~ Similarly, the ocean carbon sink over 1750-2011 of the unconstrained OSCAR is 159 ± 20 PgC, higher than the one of IPCC AR5 (Ciais et al., 2013b), 155 ± 18 PgC, in terms of mean and standard deviation. The constraints on cumulative compatible emissions mostly impacts RCP6.0 and RCP8.5, transforming the bimodal distribution of the unconstrained OSCAR into a monomodal distribution. Using this constraint, the mean of OSCAR is increased and the range decreased, reaching 163 ± 15 PgC.

Applying these constraints successfully reproduce the observed distribution, but also reduces the range in the other constraints, such as the cumulative net ocean carbon flux over 1750-2011. We note that combining these constraints leads to a tightening of the posterior distribution, thus likely introducing a bias. OSCAR could benefit from further development in this direction, following (McNeall et al., 2016; Williamson and Sansom, 2019).

3.2. Climate response

Simulations with an abrupt increase in atmospheric CO_2 (and thus in radiative forcing) are typically used to ~~diagnose~~ evaluate the climate response of complex models. We use three such experiments from CMIP6 and RCMIP with quadrupled, doubling and halving atmospheric CO_2 (*abrupt-4xCO2*, *abrupt-2xCO2* and *abrupt-0p5xCO2*). These experiments can be used to estimate the Equilibrium Climate Sensitivity (ECS) of an ESM or a model such as OSCAR (Gregory et al., 2004) and investigate how this metric is influenced by the intensity of the forcing. The~~se~~ results are shown in ~~Figure 4~~Figure .

The ECS is defined as the equilibrium temperature that results from the doubling of the preindustrial atmospheric concentration of CO_2 (Gregory et al., 2004). The ECS and its calculations have evolved with the integration of new components to climate models (Meehl et al., 2020). In regard of the computational cost of the ESMs, reaching this equilibrium takes a time long enough to use Gregory's method (Gregory et al., 2004) to calculate the ECS or alternative methods (Lurton et al., 2020; Schlund et al., 2020). The ECS using the Gregory method is actually not exactly the equilibrium climate sensitivity per se, but rather an "effective climate sensitivity" (Sherwood et al., 2020). Paleoclimate data shows that feedbacks from vegetation, biogeochemistry or dust affect the equilibrium (Friedrich et al., 2016; Rohling et al., 2012). From CMIP5 to CMIP6, ESMs have improved their treatment of the biogeochemistry and the vegetation, leading to alteration in feedbacks and aerosols fields (Meehl et al., 2020). This evolution participates in the observed changes in ECS from CMIP5 to CMIP6, attributed to cloud effects (Zelinka et al., 2020) and the pattern effect (Dong et al., 2020).

In OSCAR, there are two ways of estimating the ECS. First, because OSCAR is not process-based, the ECS is actually a parameter of the model. Since the formulation of the climate module is linear (Gasser et al., 2017; Geoffroy et al., 2013b), we also know that this value is independent of the intensity of the abrupt experiment. This parameter was calibrated on the *abrupt-4xCO2* experiment run by CMIP5 models and normalised to OSCAR's estimate of RF for a quadrupling of CO_2 (Gasser et al., 2017). Under this definition, the ECS of OSCAR follows the Gregory's method and does not account for all feedbacks of OSCAR. When using parameters from OSCAR, the climate feedbacks ~~actually~~ included in the estimated ECS depend on the CMIP5 models used for calibration. If calibrated on general circulation models (GCMs), only the so-called Charney feedbacks are included (i.e. Planck, water vapour, lapse rate, sea-ice albedo, and clouds) with the possible addition of the CO_2 physiological feedback (Sellers et al., 1996). However, when calibrated on ESMs, additional feedbacks ~~pertaining~~ relative to interactive biogeochemical cycles may be included, depending on what exact processes are implemented in a given ESM. The second way of estimating the ECS in OSCAR is to define it as the GSAT change at the end of the 1,000 years of the abrupt experiments. Here, all of the feedbacks integrated in OSCAR are accounted for, especially biogeochemical feedbacks.

Values related to these two approaches are presented in Table 2. The ECS calculated using parameters of OSCAR, hence comparable to Gregory's approach, is 2.78 ± 0.47 K when constrained, while the unconstrained one is 3.17 ± 0.63 K. ~~This~~By construction, this~~, by construction,~~ is consistent with the AR5 estimates (Collins et al., 2013), but also with more recent assessments (Gregory et al., 2020). Because we use observational constraints, these results are lower than the CMIP5

range 2.1 – 4.7 K (Andrews et al., 2012). The CMIP6 range, 1.8 – 5.6 K (Zelinka et al., 2020; Meehl et al., 2020) is even higher than the CMIP5 range. The higher values for the ECS from some CMIP6 models are significantly reduced when constraining (Nijsse et al., 2020; Bonnet et al., 2021), with some ECS estimates even lower than those shown here, such as (e.g. 1.38K with a likely range of 1.3-2.1K) than those shown here. Overall, these values provided by OSCAR remain consistent with the literature, albeit on the lower end of the range (Sherwood et al., 2020). Similarly As shown in Table 2, the Transient Climate Response (TCR) and the Transient Climate Response to Emissions (TCRE) of the unconstrained OSCAR are also consistent with the CMIP5 values (Meehl et al., 2020) and (Gillett et al., 2013), thanks to the calibration of the ECS in OSCAR. Constraining OSCAR reduces all these metrics both in value and in range, and. we We attribute this effect to the constraint on historical warming. This reduction effect is similar to what was shown recently for CMIP6 models (Tokarska et al., 2020).

The other approach to derive ECS using abrupt experiments is illustrated in Figure 4Figure . It leads in *abrupt*-2xCO2 to an unconstrained ECS of 2.74 ± 0.52 K (Table 2), reduced to 2.52 ± 0.33 K with the constraints. Overall, the ECS is remarkably consistent in terms of average, standard deviation and even skewness across the three abrupt experiments. This is due to the construction of OSCAR, with a prescribed logarithmic dependency of the radiative forcing of CO_2 to its atmospheric concentration (Lurton et al., 2020). This ECS is lower than with the first approach, because it includes several Earth system feedbacks related to short-lived species that are left free to change during the simulations, owing to the experimental protocol. In OSCAR, this is mostly explained by an increase in the atmospheric load of tropospheric aerosols (and ozone) caused by the endogenous emission of precursors through biomass burning. These feedbacks are also illustrated in Figure 4Figure . The RF resulting from the prescribed change in atmospheric CO_2 (7.42 $W.m^{-2}$ under quadrupled CO_2) is partially compensated by short-lived climate forcers. In the case of *abrupt-4xCO2*, the RF sums up to 3.46 ± 0.25 $W.m^{-2}$, because of a cooling by scattering aerosols (-0.21 ± 0.16 $W.m^{-2}$) and aerosol-cloud effects (-0.21 ± 0.15 $W.m^{-2}$), besides an additional warming from absorbing aerosols (0.13 ± 0.08 $W.m^{-2}$). Finally, from Table 2, we note that constraining reduces the parameter-based ECS by 0.44 K, while the one with all feedbacks has its ECS reduced by 0.22 K, which implies that biogeochemical feedbacks are also significantly constrained.

3.3. Carbon cycle response

The *1pctCO2* experiment, in which atmospheric CO_2 increases by +1% every year, is part of the DECK. Two variants of *1pctCO2* have been performed as part of the C4MIP exercise (Figure 5Figure). In *1pctCO2-rad*, atmospheric CO_2 only has a radiative effect on the climate system, as a preindustrial level of CO_2 is seen by the carbon cycle. In *1pctCO2-bgc*, only the carbon cycle is affected by CO_2, whereas a preindustrial CO_2 is prescribed to the climate system. The outputs of OSCAR v3.1 on these experiments are consistent with past C4MIP results (Arora et al., 2013). The global mean surface temperature responds about linearly to the exponential increase in CO_2, because of the implemented logarithmic dependency of the

Formatted: Font: (Default) +Headings (Times New Roman), 10 pt, Not Bold

radiative forcing of CO_2 to its atmospheric concentration. Carbon sinks rise in response to the increase in atmospheric CO_2, but the resulting warming dampens the sinks.

These three experiments can be used to calculate the carbon-concentration and carbon-climate feedback metrics, respectively β and γ. These metrics, defined and used in former C4MIP exercises (Friedlingstein et al., 2006; Arora et al., 2013; Arora et al., 2020), are a means to ~~diagnose~~ evaluate the model's sensitivities of the carbon stocks in the land and in the ocean to changes in atmospheric CO_2 or GSAT. Table 3~~Table 3~~ summarizes these results. As explained by (Arora et al., 2013), there are three methods to combine the three experiments to calculate the metrics: subtracting *1pctCO2-bgc* from *1pctCO2-rad* (noted R-B, hereafter), subtracting *1pctCO2* from *1pctCO2-bgc* (B-F), and subtracting *1pctCO2* from *1pctCO2-rad* (R-F). As shown in **Table** 3, ~~M~~methods R-B and B-F are almost equivalent for β, while methods R-B and R-F are almost equivalent for γ. Although LUC affects these metrics (Melnikova et al., 2021), these experiments are designed to have a constant LUC.

Table 3~~Table 3~~ shows that β under the R-F method ~~are~~ is lower than the R-B and B-F because the non-linearity of the Earth system reduces the sensitivity of land and ocean carbon to atmospheric CO_2. Similarly, γ under the R-B and R-F ~~are~~ is higher than under the B-F, but the non-linearity here is added to R-B and B-F (Arora et al., 2013). Applying our observational constraints increases the absolute values of β_{land} and γ_{land} of OSCAR, but it does not affect significantly the β_{ocean} and γ_{ocean}. The only exception is the γ_{ocean} under the method B-F. We note that the unconstrained OSCAR v3.1 is closer to the CMIP5 exercises, be it at 2x or 4x CO_2. This result can be explained with OSCAR v3.1 being calibrated on CMIP5. However, the unconstrained β_{land} is the only one to be closer to CMIP6 than to CMIP5. The cause of this difference in the β_{land} remains unclear but may come from the form of equation for the fertilization effect. The configurations of OSCAR are not only different parameters, but also different equations. Here, half of the configurations of OSCAR follow a logarithmic formulation of the fertilisation effect (Gasser et al., 2017), which may not be convex enough to properly represent a saturation effect found in many ESMs. We note that in our assessment, the land includes permafrost carbon, which was not the case in CMIP5 assessment, but the permafrost is mostly sensitive to increase in temperatures (i.e. it impacts γ_{land} but not β_{land}).

Overall, **Table 3** shows that the unconstrained carbon cycle of OSCAR v3.1 is ~~well~~ in line with CMIP exercises, particularly CMIP5. Yet, the sensitivity of the oceanic carbon stock to increase in GSAT remains too high. This bias in the ocean module could be attributed to the stratification effect introduced in v2.2 (Gasser et al., 2017). In any case, this suggests that our carbon cycle may be too optimistic, which will clearly appear in our emission-driven simulations.

3.4. Reconstruction of the historical period

The concentration- and emission-driven historical experiments (i.e. *historical* and *esm-hist*, respectively) were run with OSCAR. Their forcers differ only ~~on~~ for CO_2: the atmospheric CO_2 is prescribed in the former, whereas in the latter, fossil-

Formatted: Font: (Default) +Headings (Times New Roman), 10 pt, Not Bold

Formatted: Font: Not Bold

Formatted: Font: (Default) +Headings (Times New Roman), 10 pt, Not Bold

Formatted: Font: Bold

fuel emissions are prescribed and atmospheric CO_2 is fully interactive. In the concentrations-driven *historical*, compatible fossil-fuel emissions are back-calculated after the simulation (Jones et al., 2013; Gasser et al., 2015). Altogether, these two simulations are relatively close, as shown in ~~Figure 6~~Figure , but with noticeable differences.

Looking at the carbon- cycle variables, we observe that up to the 1940s, *esm-hist* is ~~relatively~~ similar to *historical* in terms of fossil-fuel CO_2 emissions, atmospheric CO_2 and both carbon sinks. For instance, the cumulative ocean sink over 1850-1940 are respectively 41 PgC and 35 PgC in *historical* and *esm-hist*. The difference observed afterwards can essentially be explained by the fact that the emission-driven simulation entirely misses the 1940s plateau in atmospheric CO_2. Such a miss is typical of ESMs (Bastos et al., 2016). For comparison after 1959, we use data from the Global Carbon Budget (Friedlingstein et al., 2020) whose assessed ocean carbon sink is ~~slightly~~ closer to our *historical* than to our *esm-hist*. The net carbon flux from atmosphere to land (i.e. the aggregate of the land sink, emissions from LUC, and those from permafrost) of the two historical experiments are similar from the 1980s onward. For comparison, the estimates for this average net land flux is 1.5 ± 1.1 PgC ~~/~~yr^{-1} over 2000-2009 (Friedlingstein et al., 2020) while this flux calculated by OSCAR ~~respectively~~ under *historical* and *esm-hist* are 0.88 ± 0.48 PgC ~~/~~yr^{-1} and 0.85 ± 0.56 PgC~~/~~ yr^{-1}, respectively.

Looking at the effective radiative forcings (ERF), that of CO_2 in the concentration-driven *historical* is directly deduced from the prescribed CO_2 atmospheric concentration (Meinshausen et al., 2017), but slightly higher by about 0.1 W.m^{-2} than the central value from the 5th Assessment Report (AR5) (Myhre et al., 2013). The central value from AR5 (1.82 W.m^{-2}) is calculated with reference to 1750 but becomes 1.66 W.m^{-2} when calculated with reference to 1850. This value increases to 1.70 W.m^{-2} in CMIP6 data, mostly b~~B~~ecause of changes ~~mostly~~ in the CO_2 concentration in 1850. ~~in CMIP6 data~~, ~~this value increases to 1.70 W.m^{-2}~~. With OSCAR and prescribed CO_2 emissions, the atmospheric CO_2 in *esm-hist* is higher than in *historical*, the ERF of CO_2 is 0.2 W.m^{-2} higher than in the AR5. The ERF of other greenhouse gases are consistent with (Myhre et al., 2013). For most ERF components, there is very little difference between *historical* and *esm-hist*. OSCAR's overall ability to simulate the RF of short-lived species compares well with the IPCC AR5 values. Contributions to the warming from aerosols and ozone are consistent as well, although OSCAR tends to amplify these contributions. In 2011, IPCC AR5 estimates the RF from aerosols to -1.01 ± 0.37 W.m^{-2}, while OSCAR calculates them at -1.29 ± 0.52 W.m^{-2}. Similarly, IPCC AR5 estimates the RF from tropospheric ozone in 2011 at 0.4 ± 0.2 W.m^{-2}, OSCAR being at 0.50 ± 0.05 W.m^{-2}. It may be caused by overestimated biomass burning emissions, and this will be examined more in-depth in a future analysis. Since these biases were already ~~diagnosed~~ evaluated in the description paper of OSCAR (Gasser et al., 2017), it shows that our constraining does not markedly alter these aspects of the model. Additional constraining could be introduced for separate RF components, albeit this would likely weaken the efficiency of other existing constraints.

Looking at climate variables, the increase in GSAT in both historical experiments are consistent with the Special Report on Global Warming of 1.5C (IPCC, 2018) and with the historical reconstruction by (Cowtan and Way, 2013). During the choice of constraints (sections 2.3 and 3.1, appendix A), we observed that constraints on temperatures impact much more our results than the other type of constraints. Even while the set of constraints is expanded, constraints on temperature have a lasting influence over all outputs. The *esm-hist* simulation shows a higher GSAT and appears to be further away from the

Formatted: Superscript

Formatted: Superscript

Formatted: Superscript

observations. This is mostly the result of the higher atmospheric CO_2 seen earlier, and it suggests a different set of constraining weights could be used for the emission-driven runs. We choose not to, for the sake of consistency. Comparing the effective radiative forcing (ERF) of OSCAR to the one of the IPCC AR5 (Myhre et al., 2013), we note differences caused by volcanic eruptions. Beyond the update of the time-series of volcanic activity itself, OSCAR make use of a warming efficacy of 0.6 for stratospheric volcanic aerosols (Gasser et al., 2017; Gregory et al., 2016). Nevertheless, IPCC AR5 estimates the ERF at 2.3 ± 1.0 $W.m^{-2}$ while OSCAR calculates respectively under *historical* and *esm-hist* 2.24 ± 0.48 $W.m^{-2}$ and 2.34 ± 0.50 $W.m^{-2}$. Finally, the total ocean heat content is well reconstructed, although the range of OSCAR is larger than the observed one (von Schuckmann et al., 2020), suggesting this could also be considered a potential constraint for the model in future work.

3.5. Attributions

DAMIP (Gillett et al., 2016) designed a number of experiments meant to attribute the observed climate change to anthropogenic and natural factors. Since OSCAR does not feature any internal variability, it cannot contribute to the "detection" part of DAMIP. However, with more than 1000 Monte Carlo elements, OSCAR is fully capable of carrying out the "attribution" part. To ~~do~~ achieve this attribution, DAMIP relies on experiments that follow the *historical* one, but in which only one forcing is turned on. Conversely, a number of other MIPs introduced attribution experiments in which all forcings but the ones studied are turned on. However, neither of these approaches explicitly considers the non-linearities of the system. Other more robust methods of attribution to forcings exist (Trudinger and Enting, 2005) and have been used with OSCAR in the past (Gasser, 2014; Li et al., 2016; Fu et al., 2020; Ciais et al., 2013a). Here, we focus on results made possible with the CMIP6 experiments, ~~that~~ which are presented in Table 4.

In the *historical* experiment, we find a change in GSAT of 0.98 ± 0.17 K in 2006-2015 with regard to 1850-1900, which is in line with observations because of our constraining setup (Section 2.3). Natural forcings caused only ~0.03 K of this total, of which ~0.02 and ~0.01 were respectively caused by solar and volcanic activity. Note that our volcano-related forcing is defined against an average and constant volcanic activity during the preindustrial period. ~~, which explains the~~ This is why the volcanic activity contributes only a positive~~to~~ ~0.01K ~~(slightly) positive response caused by this forcing~~ over the recent past where no major volcanic eruption happened. In the IPCC terminology, our results lead to the conclusion that it is *extremely unlikely* (i.e. likelihood <1%) that natural factors alone are causing the current observed climate change. This is of course consistent with the IPCC conclusions (Eyring et al., 2021; Gillett et al., 2021). Nevertheless, we note that our constraining reduces the uncertainty range of all simulations, including those driven only by natural forcings. For the simulations under natural forcings, the range from the constrained OSCAR is smaller than the ones from (Gillett et al., 2021), which may suggest an over-constraining. It may be solved by using different methods for constraining climate simulations (Nicholls et al., 2021; Williamson and Sansom, 2019).

Since DAMIP did not include an experiment in which only natural forcings would be turned off, we cannot conclude as to the ~~(Gillett et al., 2021)~~ complementary probability of observed climate change being caused only by anthropogenic factors (Gillett et al., 2021). Attribution to groups of anthropogenic forcings is possible, however. We find that 1.25 ± 0.11 K, about 128 % of the recent warming, was caused by well-mixed greenhouse gases (WMGHGs) and -0.26 ± 0.22 K (-27 %) was by near-term climate forcers (NTCFs). For comparison, the 90% confidence interval of CMIP6 over 2010-2019 instead of 2006-2015 are 1.16 to 1.95 K for WMGHGs and -0.73 to -0.14 K for NTCFs (Gillett et al., 2021). Another contribution of -0.03 ± 0.03 K (-3%) is due to land-use change. We highlight that observational constraints affects these contributions, as shown by~~in~~ (Ribes et al., 2021), whose~~.~~ ~~The~~ central estimate contributions ~~corresponding to their central estimate~~ over 2010-2019 are 116% for WMGHGs and -32% for NTCFs and land-use change. It follows that the constrained results of OSCAR v3.1 are consistent with (Gillett et al., 2021; Ribes et al., 2021).

Considering the other experiments, we observe that the DAMIP experiment (*hist-aer*) and the AerChemMIP one (*hist-piNTCF*) led to very similar estimates of the contribution of NTCFs (Table 4), which highlights that this part of our model behaves in a ~~fairly~~ linear fashion. Going further in isolating individual forcings, we also estimate that CO_2 caused 0.74 ± 0.06 K, chlorofluorocarbons and hydro-chlorofluorocarbons (i.e., CFCs and HCFCs) caused 0.13 ± 0.02 K, stratospheric O_3 caused -0.03 ± 0.03 K, and all aerosols together caused -0.33 ± 0.21 K (including direct and indirect effects). We point out that details on CH_4, N_2O or tropospheric ozone cannot be provided, because of the lack of relevant CMIP6 experiments.

The extent to which this attribution to specific forcings is comparable to existing studies remains ~~debatable~~unclear. One notable limitation of OSCAR, in this respect, is that the model's climate response is not forcing-dependent. The use of effective radiative forcing is supposed to ensure that the temperature response to CO_2 and non-CO_2 forcings is similar, at least for the long-term steady-state (Myhre et al., 2013). However, recent work has pointed out that the response may strongly depend on the forcing agent (Marvel et al., 2016), thus casting a degree of doubt on our attribution results. More work to integrate such differentiated responses in reduced-complexity models is warranted.

3.6. Scenarios of climate change

ScenarioMIP (O'Neill et al., 2016) cho~~o~~se eight particular SSPs taken from the SSP scenario database (Riahi et al., 2017) to cover a range of socio-economic assumptions and climate targets~~., and then harmonised them to~~ After harmonization, these SSPs bec~~a~~~~o~~me the default CMIP6 scenarios to be run by ESMs (Gidden et al., 2019). ScenarioMIP mostly required concentration-driven simulations up to year 2100 or 2300~~, however, which was complemented i~~. In RCMIP, this was~~ey were~~ complemented by extending all scenarios up to 2500 and systematically running emission-driven simulations in addition (Nicholls et al., 2020). ~~Figure 7~~ Figure displays projections of key global variables of the Earth system following these scenarios, and Table 5 focuses on projected GSAT changes.

The climate target dimension of the SSP scenarios is defined similarly to the RCPs' as the total RF targeted in 2100 (van Vuuren et al., 2011). Table 5 shows that this targeted RF is overall within the 1-σ uncertainty range of all our concentrations-driven projections. In the cases with notable differences, such as *ssp460*, the actual RF reached by the reduced-complexity

Formatted: Font: Not Bold

model MAGICC (IIASA, 2018b) for this scenario is 5.29 $W.m^{-2}$, which is then in the range of OSCAR. ~~Because~~ Although MAGICC ~~has been~~was used for the design of these scenarios, this result~~it~~ demonstrates that we remain consistent with the intended RF of the scenarios. Emission-driven SSPs show lower RF than their concentration-driven counterpart, which can be attributed to a low bias in the atmospheric CO_2 that is especially visible in high CO_2 scenarios. This bias is a result of our constraining approach that favoured configurations with strong CO_2-fertilization (as also seen with the C4MIP results, section 3.3). Under high CO_2 scenarios, this bias is likely worsened by our exclusion procedure during the post-processing, as very high CO_2 tends to make the model more unstable. The very low uncertainty range we obtain for projected atmospheric CO_2 in emission-driven simulations ~~seems~~ is over-confident. ~~However, w~~We note that the constraints were derived using concentration-driven simulations (that are the focus of CMIP6), and so they may not apply properly to emission-driven simulations.

The constraining approach contributes to having the increases in GSAT for concentration-driven experiments shown in Table 5 ~~for concentration-driven experiments~~ to be lower than the CMIP6 models we could compare our results to~~.~~. The uncertainty range simulated by OSCAR is also much lower, again owing to our constraining approach. With a relative uncertainty in GSAT change in 2500 of ±13% under the warmest scenario (SSP5-8.5), ~~one may wonder whether~~ these projections are likely to be over-constrained~~, extending the constrain in observed GSAT throughout the scenario~~. This stems from our constraining of the climate response, as also shown by the relatively small uncertainty range in ECS in the idealised abrupt CO_2 experiments. Further developing that module by adding one or two key parameters (Geoffroy et al., 2013a; Bloch-Johnson et al., 2015) would provide more degrees of freedom and likely release part of the constraint. When projecting temperature change in an emission-driven mode, ~~however,~~ the uncertainty range is larger~~,~~ because of the additional uncertainty related to the biogeochemical cycles.

The CMIP6 values are here computed from CMIP6 time series. However, some CMIP6 models exhibit higher warmings than in previous assessments, and observations can be used to constrain the future warming (Tokarska et al., 2020). Using their table S4, the warming in 2081-2100 with reference to 1995-2014 under SSP5-8.5 for the constrained CMIP6 models is 3.44 ± 0.67 K and 3.11 ± 0.36 K for OSCAR v3.1 constrained. For SSP1-2.6, the values are respectively 0.94 ± 0.30 K and 0.76 ± 0.17 K. Thus, the observational constraints that we have used contribute to explain the differences to the raw CMIP6 data. Nevertheless, it remains that the climate module of OSCAR v3.1 could still be improved.

3.7. Zero Emissions Commitment

ZECMIP aims at investigating the zero-emission commitment (ZEC), that is the additional warming that follows a cessation of anthropogenic CO_2 emissions (Jones et al., 2019). Two categories of experiments were performed. The first one (called branched experiments) is a variation of the emission-driven *1pctCO2*, in which emissions cease once they reach 750 PgC, 1000 PgC or 2000 PgC of cumulative value. These distinct levels of cumulative emissions are meant to evaluate the state dependency of ZEC. The second category consists in three bell-shaped emission pathways whose cumulative emissions

are the same as in the branched experiments. This was proposed by ZECMIP to evaluate the dependency of the ZEC on CO_2 emission rate, as the emission rate at the time of cessation is near zero in these bell experiments ~~(whereas it is,~~ while very high in the branched ones~~)~~.

~~Figure 8~~ Figure shows the time series of the ZEC in both sets of experiments. In the branched experiments, the abrupt cessation of CO_2 emissions triggers an abrupt increase of temperature change, followed by a decrease. Conversely, ~~in the bell experiments,~~ since the cessation is smoother in the bell experiments, no abrupt response is visible on the very short term. After this period, the shape of the evolutions of the ZEC in branched experiments is similar to the shape in bell experiments. We attribute this effect to the abrupt cessation of emissions in the branched experiments, causing biomass burning and aerosol lifetime feedbacks ~~(the same that affect the ECS)~~ whose response to temperature change happens within the same year. These feedbacks explain why the ZEC in branched experiments is systematically lower than the ZEC in bell experiments.

~~Figure 8~~ Figure also shows that the ZEC for a cumulative emission of 2000 PgC is much higher than in the two other cases, highlighting a strong non-linearity in the model. We attribute this process to the permafrost response, in ~~complete~~ agreement with our previous work (Gasser et al., 2018). Once the branching year has been reached, anthropogenic emissions become zero, while natural systems such as the permafrost keep emitting. ~~Under higher warming, the existing warming at cessation of emissions comes with a legacy, and permafrost contributes to the non-linearity of this legacy.~~ Among the models that contributed to ZECMIP (MacDougall et al., 2020), CESM2, NorESM2-LM and UVic ESCM 2.10 were the only ones to model permafrost, with only the later one that provided data over the three branched experiments. As shown in Figure 6 of (MacDougall et al., 2020), UVic ESCM 2.10 is the model with the strongest evolution of the ZEC with cumulative emissions. This similar effect of permafrost on ZEC in OSCAR v3.1 and UVic ESCM 2.10 calls for more contributions of models with permafrost to the ZECMIP exercise and future similar projects.

As illustrated in Table 6, OSCAR v3.1 estimates a ZEC (in the reference case of the *esm-1pct-brch-1000PgC* experiment) that is within the range of ZECMIP (MacDougall et al., 2020). The evolutions of OSCAR in this experiment are ~~similar~~ comparable to those of the Earth system models of intermediate complexity that contributed to the original ZECMIP.

3.8. Behaviour of OSCARv3.1

The focus of this paper is to ~~diagnose~~ evaluate this version of OSCAR introduced in (Gasser et al., 2020a), and used with the same exclusion and constraining approach used for RCMIP phase 2 (Nicholls et al., 2021). As explained in section 2.2, many experiments have been run through OSCAR v3.1, and sections 3.1 to 3.7 have used only the experiments that would allow clear comparison with ESMs and therefore ~~diagnosis~~evaluation. In the Appendix B, additional results are shown, further illustrating the behaviour of OSCAR v3.1 ~~on~~ under experiments that examine carbon geoengineering (section ~~A~~B.1), solar geoengineering (section ~~A~~B.2), land-use (section ~~A~~B.3), NTCFs (section ~~A~~B.4) and a comparison of RCPs against~~to~~ SSPs (section ~~A~~B.5). These additional experiments were not fully considered in the ~~diagnosis~~ evaluation part of this study, typically because of the lack of published papers doing the same with fully fledged ESMs or because of non-existent

~~diagnostic~~ evaluation metrics. These simulations can nevertheless provide valuable insights into the behaviour of OSCAR, potentially helping understand past or even future contributions to community exercises such as CDRMIP or RCMIP.

4. Concluding remarks

In this study, we ~~have~~ present~~ed~~ the setup used with OSCAR v3.1 to run 75 CMIP6 and 24 additional experiments from RCMIP. We ~~have used~~use the primary results of these simulations to discuss the overall behaviour and performance of our model, comparing our results to those of state-of-the-art complex models whenever possible. We present below a brief summary of the model's main limitations.

First, the model tends to be unstable under high CO_2 and high warming scenarios. This comes mostly from the ocean carbon cycle module whose stability is not ensured under our chosen differential system solving scheme, which is also worsened by the stratification feedback that was introduced in v2.2 (Gasser et al., 2017). This pleads for a revamp of this module.

Second, despite a clear improvement of the land carbon cycle module in v3.1 (Gasser et al., 2020a), its unconstrained transient response remains wider than the ranges from CMIP5 or CMIP6 models, which makes the constraining step a strong requirement of any simulation with OSCAR. In its current state, the constraining step appears to favour parameterizations with a strong CO_2 fertilization effect. The extent to which this is caused by structural modelling choices is unclear. Consequently, the land carbon cycle also exhibits a sensitivity to climate change that is too low compared to complex models, mostly those without permafrost, thus calling for an improved calibration. ~~A potential track would be to account for correlations between parameters within the prior distribution of parameters (i.e., for instance when drawning from the Monte Carlo).~~

Third, the constrained climate module shows a relatively low ECS and a rather narrow uncertainty range. Introducing extra parameters for the heat uptake feedback (Geoffroy et al., 2013a) and possibly non-linear Charney feedbacks (Bloch-Johnson et al., 2015) would likely help to gain flexibility during the constraining. This third point is the reason behind most of the difference between OSCAR and CMIP6 temperature projections shown in Table 5.

Fourth, although most of the non-CO_2 species are reasonably simulated, the effects of tropospheric ozone and total aerosols tend to be overestimated. The whole aerosol module behaves rather linearly, and it exhibits a climate feedback whose intensity should be better constrained against existing simulations with complex ESMs. OSCAR would indeed benefit from further work on short-lived species, although this could prove a challenging endeavour given the aggregated formulation of the model and the uncertainties.

Finally, we have illustrated how observational constraints can be used to inform projections, how it may affect the results, such as the strong decrease of uncertainties in projections. Given the growing importance of these constraints (Tokarska et al., 2020; Nicholls et al., 2021), this calls for investigating computationally efficient and physically sensible

ways of doing so with OSCAR. Investigating and controlling the bias introduced in these steps may increase the confidence in the model's results (McNeall et al., 2016; Williamson and Sansom, 2019).

In spite of those limitations, we have demonstrated that OSCAR behaves as one would expect from an Earth system model. Applying our two post-processing steps (exclusion and constraining) overcomes some of the model's limitations, and the resulting quantitative behaviour of OSCAR ~~remains largely satisfactory~~is thus improved. In several cases, we have also shown that OSCAR differs from complex models, due to ~~the~~ features ~~of OSCAR~~ that are not yet ~~in~~ part of most complex models, such as ~~fully interactive atmospheric chemistry that would allow CH_4 and N_2O to be emission-driven, and~~ endogenous simulation of CH_4 emissions from wetlands, CO_2 and CH_4 emissions from permafrost, and emissions from biomass burning. Therefore, ~~some of~~ the results presented here have scientific interests that go beyond the pure model evaluation perspective. ~~These valuable insights for other projects will be presented in separate studies~~To this intent, ~~but~~ many outputs from the simulations presented here are already publicly available as part of the RCMIP exercise (Nicholls et al., 2021). More outputs can be requested from the authors. Finally, this study will be the basis for a more systematic assessment of the model's performance, as we will use the standardised CMIP6 and RCMIP simulations to ~~diagnose~~ evaluate future versions of OSCAR and to compare them with older versions. This will provide the wider community with a benchmark of the model, hopefully spreading interest in this open-source compact Earth system model.

~~Appendix A: Behaviour of OSCAR~~

~~A.1. Carbon geoengineering~~

~~A.1.1. Idealized experiments~~

~~Experiments of CDRMIP are designed to investigate the consequences of carbon dioxide removal for the Earth system (Keller et al., 2018a). In *1pctCO2-cdr*, the atmospheric CO_2 increases by 1% every year (just like *1pctCO2*), but after 140 years, the atmospheric CO_2 decreases following a pathway at the same rate than the ramp-up period. Once CO_2 has returned to its preindustrial state, the experiment is extended over 1000 years. As shown in Figure A. 1, the GSAT reaches 3.68 ± 0.39 K at the end of the ramp-up forcing, and it goes back to 0.85 ± 0.22 K at the end of the ramp-down forcing. For all variables, such as the CH_4 emissions from wetlands, removing CO_2 from the atmosphere during ramp-down effectively reduces the perturbation in the variable that was induced by the ramp-up, albeit within a different time frame that is typical of a dynamic hysteresis (Boucher et al., 2012). Even the permafrost carbon stock slowly reconstitutes itself once the global temperature change is sufficiently reduced. However, the whole Earth system is not fully recovered as soon as the preindustrial level of atmospheric CO_2 is reached. To return within 10% of the maximum perturbation at the end of the CO_2 ramp-up, it takes GSAT an average 110 extra years, and the land carbon stock an average 26 years. At the end of the 1000-year extension, the oceanic carbon stock remains at about 19% of its maximum perturbation.~~

Other CDRMIP experiments based on pulses of carbon emission or removal in an emission-driven configuration were performed to evaluate the response of the Earth system to CDR. These experiments are used to calculate the Absolute Global Warming and Temperature Potentials (AGWPs and AGTPs) of CO_2, which serves to establish the Global Warming and Temperature Potentials (GWPs and GTPs) of other greenhouse gases (Myhre et al., 2013). In *esm-pi-CO2pulse*, a 100 PgC pulse is emitted from the preindustrial environmental condition in 1860, whereas 100 PgC are removed in *esm-pi-cdr-pulse*. In *esm-yr2010CO2-CO2pulse*, the 100 PgC pulse is applied in 2015 but under 2010 environmental conditions, whereas these 100 PgC are removed at the same date in *esm-yr2010CO2-cdr-pulse*. We calculate time series of AGWPs and AGTPs under these experiments (Figure A. 2) and note how close they are. We pinpoint that, just like the other experiments, we are calculating these potentials with the interactive permafrost of OSCAR. The larger source of differences lies in the background: under preindustrial environmental conditions, emission pulses have a stronger AGWP or AGTP over the short term, but this is the opposite over the longer term. Over the short term, this is due to the logarithmic expression of the CO_2 radiative forcing that is less saturated under preindustrial conditions. Over the long term, this is due to the deterioration of the carbon sink capacities under current conditions (Raupach et al., 2014). Similar reasons explain why a pulse of carbon removal cools the atmosphere slightly more over the short term than a pulse of emission warms it, but less over the long term. Our results cannot be compared to the final CDRMIP results yet, for they are unpublished, but they are consistent with those obtained with a model of intermediate complexity (Zickfeld et al., 2021).

A.1.2. Alternative scenarios

The C4MIP (Jones et al., 2016) experiments *ssp534-over-bgc* and *ssp585-bgc* differ from *ssp534-over* and *ssp585* in that the prescribed CO_2 does not affect the total radiative forcing , thus causing a lower change in GSAT and maintaining a relatively high carbon sinks efficiency. Figure A. 3 shows both carbon sinks under the variants and the base scenarios. Note that the *-bgc* experiments stem from a different historical simulation (*hist-bgc*). Under the high warming scenarios *ssp585*, climate change reduces the oceanic carbon sink by 1.93 ± 0.69 $PgC.yr^{-1}$ and the net land carbon flux by 4.31 ± 1.93 $PgC.yr^{-1}$ in 2100. Under the overshoot scenario *ssp534-over*, this difference is lower, owing to its declining atmospheric CO_2. Removing the impact of climate change on the carbon cycle increases the land carbon stock by 269 ± 52 PgC in *ssp534-over*, but by 501 ± 117 PgC in *ssp585* in 2100, due to the higher warming in the latter case. We note that the permafrost carbon stock drives most of the changes, because if permafrost is ignored in the *bgc* variant, these changes are reduced to 57 ± 32 PgC and 131 ± 77 PgC in *ssp534-over* and *ssp585* respectively.

A.2. Solar geoengineering

A.2.1. Idealized experiments

Experiments of GeoMIP (Kravitz et al., 2015) are designed to investigate the geoengineering techniques of Solar Radiation Management (SRM). Although OSCAR is not suited for all GeoMIP experiments, as it lacks any spatially resolved process, a few simulations remained accessible to our model. We run experiments *G1* and *G2*: G1 essentially follows *abrupt-4xCO2*, albeit with a changed incoming solar radiation that compensates for the radiative forcing caused by the increasing atmospheric CO_2. For *G2*, an identical principle is applied but using *1pctCO2* as a basis. As explained by (Kravitz et al., 2011), the change in solar radiation compensates solely for the radiative forcing of CO_2. However, it does not compensate for other radiative effects introduced by biogeochemical feedbacks, such as the fertilization by CO_2, affecting the carbon cycle, thus changing biomass burning emissions. Figure A. 4 shows that offsetting the CO_2 radiative forcing with a change in solar activity effectively compensates the change in GSAT. However, we simulate that the GSAT decreases in *G1* and *G2* to reach -0.08 ± 0.20 K and -0.07 ± 0.20 K, respectively, at the end of simulations. The compensation of the sole radiative forcing of CO_2 does not balance other feedbacks. There remains an additional radiative forcing, mostly due to changes in aerosols (as also shown in Figure 4), which results in this relatively small cooling in *G1* and *G2*. We estimate that in OSCAR about half of this effect is caused by the vegetation being fertilized by CO_2 and fuelling increased natural biomass burning emissions, and the remaining half is caused by the direct impact of GSAT on the atmospheric lifetime of aerosols (not shown). We note that the latter effect could be poorly estimated, in these specific experiments, as OSCAR's formulation for the lifetime of aerosols depends only on GSAT and not on the precipitation intensity.

Indeed, global precipitation does not respond in a similar way, because changes in atmospheric CO_2 and solar radiation have a different impact of the hydrological cycle (Andrews et al., 2010). In spite of a fully compensated GSAT change, global precipitation is significantly reduced in *G1* and *G2*, showing that such SRM technique does not entirely negate climate change. This demonstrates that OSCAR is capable of reproducing this well-established effect of this SRM technique (Boucher et al., 2013b). One added value of having a fully coupled ESM run these GeoMIP experiments is that we can also provide an estimate of the impact of the SRM technique on the carbon cycle. Figure A. 4 also shows that the land and ocean carbon stocks are increased in *G1* and *G2*, respectively by about 33% and 20% at the end of the simulations, owing to the loss of carbon sink efficiency that is avoided by maintaining the temperature to its preindustrial level.

A.2.2. Alternative scenarios

In addition to the few idealized experiments of GeoMIP (Kravitz et al., 2015) that are accessible to OSCAR, one scenario variant focusing on SRM was also feasible. The *G6solar* experiment stems from *ssp585*, but the solar constant is changed from 2020 onwards to compensate the radiative forcing of *ssp585* and match the one of *ssp245*. As shown in Figure A. 5, although the GSAT of *G6solar* is brought to a level comparable to *ssp245*, difference remains. The change in solar

constant is calculated ex-ante as the difference from the radiative forcing of *ssp245* to *ssp585*, which by construction excludes feedbacks caused by this change and (as with *G1* and *G2*) does not fully cancel the change in global precipitation. Consequently, the carbon stocks still increase in *G6solar*, even more than in *ssp585* thanks to the lower GSAT and despite lower global precipitation.

A.3. Land-use

A.3.1. Alternative historicals

LUMIP consists of experiments specifically focusing on land-use activities, and most of them are run by the Earth system models in a so-called "offline" fashion (Lawrence et al., 2016). It means that a reconstruction of past climate variables GSWP3 (Lawrence et al., 2016; van den Hurk et al., 2016) is prescribed to the model, so that the land module is actually decoupled from the rest of the model. Despite its simplicity, OSCAR has an added-value in running those simulations, as it embeds a book-keeping module that endogenously estimates CO_2 emissions from land-use and land-cover change. The main land carbon fluxes and stocks simulated under the reference experiment (dubbed *land-hist*) are shown in Figure A. 6, along with three sets of sensitivity experiments described hereafter. The results are similar to those obtained recently with the same version of the model but with slightly differing forcings and a different constraint (Gasser et al., 2020a). The simulated land carbon stock decreases up to the 1970s, because of land-use activities emitting more CO_2 than the sink absorbs thanks to CO_2 fertilization and other factors. The carbon stock of 2010 is higher than the one of 1850 by only 1 ± 42 PgC. For comparison, the GCB 2020 provides for 1850-2014 a net budget for the land sink and CO_2 emissions from LUC of -5 ± 90 PgC (Friedlingstein et al., 2020).

The experiments *land-cCO2* and *land-cClim* are used to disentangle the contribution of CO_2 fertilization and changing climate on the land carbon cycle. In *land-cCO2*, the atmospheric CO_2 is constant and set to preindustrial value. In *land-cClim*, the climate drivers loop over the year 1901-1920 of the data set, thus simulating a preindustrial climate. Figure A. 6 shows the differences; for example, *land-hist* – *land-cCO2* illustrates the effect of atmospheric CO_2 on the variables of interest. Thanks to these experiments, we show that CO_2 is the main driver of the land sink in OSCAR, driving most of the trend, with climate bringing a significant interannual variability but virtually no trend, except over the recent past. In 2010, climate caused a small difference of -10 ± 10 PgC in total land carbon stock, while CO_2 did one of 141 ± 42 PgC. This has to be balanced with the results of the C4MIP idealized experiments, however, where we saw OSCAR is less sensitive to climate change than CMIP5 models. Additionally, we see that the effect of climate and CO_2 on land-use and land-cover change emissions is minor, which is consistent with the fact that they are firstly determined by preindustrial carbon densities (Gasser et al., 2020a; Gasser and Ciais, 2013b).

A second set of experiments is meant to Investigate the impact of land-use practices. Land-cover change contributed -152 ± 44 PgC to the 2010 change in land carbon stock since 1850, which corresponds to most of the total land-use and land-

cover change emissions. Notably, it also reduced the land sink – an effect called the loss of additional sink capacity that has been diagnosed and quantified with OSCAR in the past (Gasser et al., 2020a; Le Quéré et al., 2018a; Gasser and Ciais, 2013b; Friedlingstein et al., 2020). Shifting cultivation (i.e. rapidly rotating land-use change between agriculture and natural ecosystems) had a relatively low impact on CO_2 emissions, leading to a change in land carbon stock of -8 ± 2 PgC at the end of the simulation in 2010. Similarly, wood harvest (in woody ecosystems that do not see land-cover change) had an overall impact of -16 ± 4 PgC. Both shifting cultivation and wood harvest have no impact at all on the land sink, by construction of their formulation in OSCAR (Gasser et al., 2020a). Finally, the effect of having cropland-specific parameters in the model is isolated thanks to the *land-crop-grass* experiment, in which new croplands are treated as grasslands. Having grasslands instead of croplands increases both the land sink and the CO_2 emissions from land-use and land-cover change, resulting in a land carbon stock higher by 31 ± 26 PgC. All these values are entirely in line with an existing assessment of those land-use practices in which an earlier version of OSCAR took part (Arneth et al., 2017).

The third set of experiments relates to varying input data sets of land-use and land-cover change drivers. Two of these (*land-hist-altLu1* and *land-hist-altLu2*) relied on the two variations of the main LUH2 data set known as the "High" and "Low" variants (respectively) (Hurtt et al., 2020). We find that the so-called low variant leads to slightly higher land-use and land-cover change emissions amounting to a land carbon stock lower by 8 ± 2 PgC over the whole period. The high variant produces slightly lower total emissions, leading to a land carbon stock higher by 17 ± 5 PgC. Neither variant has a significant impact on the land sink. According to the description of these two variations (Hurtt et al., 2020), they differ from the default data set mostly in the harvest of biomass, and are very similar from 1920 onwards. The last LUMIP experiment run with OSCAR is one that uses the primary data set but an alternative starting year (*land-hist-altStartYear*). This required making an additional spin-up of the model under the environmental conditions and land cover of year 1700. Compared to the reference experiment, we find a slightly higher land sink after 1850 that decreases through time, owing to the ecosystems not being at steady state at that date. Similarly, emissions are slightly higher but the difference to the reference case tends towards zero as the legacy of land-use and land-cover change prior to 1850 fades away. The land carbon stock in 2010 is dominated by the increased land sink and amounts to a slight increase of -17 ± 13 PgC in the land. Comparing the latter value with the total change in land carbon in the reference experiment suggests starting simulations in 1850 instead of 1700 or 1750 introduces a non-negligible bias in the CMIP6 exercise.

A.3.2. Alternative scenarios

LUMIP introduced variants of regular scenarios in which alternative land-use and land-cover change drivers coming from another scenario are prescribed (Lawrence et al., 2016), some of which being used in CDRMIP to assess afforestation (Keller et al., 2018a). Two such experiments are the pessimistic *ssp585* and *ssp370* combined with the land-use activities of the optimistic *ssp126* (named *ssp585-ssp126Lu* and *ssp370-ssp126Lu*, respectively). A third experiment consists in using the land-use of *ssp370* but under *ssp126* (named *ssp126-ssp370Lu*). Comparisons of these experiments with their regular

counterparts are shown in Figure A. 7. As expected, changing the land-use scenario roughly replaces one SSP's land-use emissions by another's, albeit with slight differences in the later stage of the simulations (i.e. after 2050) when atmospheric CO_2 and climate are significantly different from the reference scenario's, which has an impact in OSCAR because of transiently changing land carbon densities. The effect on the land carbon sink is also quantified, showing that sink capacity can be preserved by conserving natural ecosystems, although it remains a relatively small effect in absolute value. We note that the ability of properly isolating both effects (on land-use emissions and on the sink) is a specific feature of OSCAR that stems from the formulation of its land carbon cycle (Gasser et al., 2020a; Gasser and Ciais, 2013b), and we do not expect many complex ESMs to be able to provide such a partitioning. The overall effect on land carbon stock change in 2100 is 48 ± 15 PgC, 76 ± 28 PgC and -65 ± 23 PgC, in the *ssp585-ssp126Lu*, *ssp370-ssp126Lu* and *ssp126-ssp370Lu* scenarios respectively. While the land carbon stocks are affected, the change in land cover also affects the planetary albedo. The radiative forcing from albedo of land cover change is exchanged between *ssp126* and *ssp370*, but changes remain below 0.1 $W.m^{-2}$. The net combined effect on projected temperature cannot be estimated because these experiments are concentration-driven.

A.4. Alternative scenarios for NTCFs

The *ssp370-lowNTCF* scenario is a variant of the *ssp370* differing by its lower emission of short-lived pollutants affecting the RF of NTCFs. As illustrated in Figure A. 8, the variant leads to an equivalent warming, although with very slightly less cooling from NTCFs. This almost negligible effect on global temperature is the result of two large but compensating effects that manifest the most between 2050 and 2100. The lower emission of warming NTCFs leads to absorbing aerosols (i.e. BC) warming less by 0.21 ± 0.11 W m^{-2} and tropospheric ozone warming less by 0.21 ± 0.03 W m^{-2} in 2100. Conversely, it also leads to scattering aerosols cooling less by 0.33 ± 0.12 W m^{-2} and the indirect aerosol effects cooling less by 0.26 ± 0.13 W m^{-2} at the same date. This results in a slight increase of the total radiative forcing of 0.15 ± 0.20 W m^{-2} and a GSAT change of only 0.07 ± 0.11 K. However, the difference in forcing agents between the two scenarios leads to a change in global precipitation that reaches 15 ± 11 mm yr^{-1} in 2100. The change in precipitation is consistent with our results for the GeoMIP experiments and what we know of the global water cycle (Shine et al., 2015).

A.5. Comparison of two generations of scenarios

Initially, the SSPs scenarios were designed to reach the RF of RCPs in 2100, to provide a common grid for reading and comparing all the SSPs scenarios. Hence, the same four RF targets chosen in CMIP5 with the RCPs (2.6 $W.m^{-2}$, 4.5 $W.m^{-2}$, 6.0 $W.m^{-2}$, 8.5 $W.m^{-2}$) have also been chosen in CMIP6 with four out of the eight SSPs used. Yet, CMIP6 ESMs did not run RCPs, because these scenarios are not part of the CMIP6 experiments. Therefore, the difference between RCPs projections in CMIP5 and SSPs projections in CMIP6 under the same RF targets are due to both a change in the generation of ESMs and

~~a change in scenarios. In Figure A. 9, we represent both RCPs and SSPs under the same version of OSCAR, showing the difference due to the sole change in scenarios. These scenarios use different drivers, as illustrated with the atmospheric CO_2 prescribed to these concentration-driven experiments, usually with higher CO_2 concentrations in the CMIP6 version. Except for the 8.5 target, the RF tends to be also higher in the CMIP6 version, compared to the CMIP5 version, meaning changes in other drivers are not enough to balance the CO_2 increase. While the 2.6 $W.m^{-2}$ and 8.5 $W.m^{-2}$ targets are reached in 2100, the 4.5 $W.m^{-2}$ and 6.0 $W.m^{-2}$ are not. However, our results can be compared to those of MAGICC in these two cases~~ (IIASA, 2018a)~~, and both reduced-complexity models are consistent. Because of the similar RF targets, GSAT are relatively similar over the 21st century, but RCPs and SSPs tend to dissociate later. In 2300, moving from RCPs to SSPs changes GSAT by 18 ± 8%, 9 ± 3%, 5 ± 2% and 6 ± 1% in the four tested scenarios, respectively. Differences in other key variables such as the carbon sinks logically respond to these differences in atmospheric CO_2 and global temperature change, as also shown in Figure A. 9.~~

Acknowledgements

TG and YQ acknowledge support from the European Research Council Synergy project "Imbalance-P" (grant ERC-2013-SyG-610028). This work is also part of the European Union's Horizon 2020 research and innovation project "CONSTRAIN" (grant #820829). This work was also partially funded by the Austrian Science Fund FWF project "ERM" (grant P31796-N29). The simulations were performed on the *ebro* cluster at IIASA. We also thank Zebedee Nicholls for providing us with the time series of global surface air temperature of CMIP6.

Author contribution

TG developed OSCAR. YQ processed the input data and set up the simulations. YQ and TG decided on the post-processing (exclusion and constraining). YQ executed the post-processing. YQ produced all the figures. YQ and TG drafted the manuscript. All authors contributed to the final analysis and to the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Code and data availability

The source code of OSCAR is available at https://github.com/tgasser/OSCAR. The code used for computation, post-treatment and plots is available at https://github.com/yquilcaille/scripts_OSCAR_CMIP6. Results for some of the experiments performed are available on the repository of RCMIP phase 2 (https://doi.org/10.5281/zenodo.4269711). Additional data ~~and scripts~~ are available upon request to the corresponding author.

References

Andrews, T., Gregory, J. M., Webb, M. J., and Taylor, K. E.: Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models, Geophysical Research Letters, 39, 10.1029/2012GL051607, 2012.

Andrews, T., Forster, P. M., Boucher, O., Bellouin, N., and Jones, A.: Precipitation, radiative forcing and global temperature change, Geophysical Research Letters, 37, 10.1029/2010GL043991, 2010.

Arneth, A., Sitch, S., Pongratz, J., Stocker, B. D., Ciais, P., Poulter, B., Bayer, A. D., Bondeau, A., Calle, L., Chini, L. P., Gasser, T., Fader, M., Friedlingstein, P., Kato, E., Li, W., Lindeskog, M., Nabel, J. E. M. S., Pugh, T. A. M., Robertson, E., Viovy, N., Yue, C., and Zaehle, S.: Historical carbon dioxide emissions caused by land-use changes are possibly larger than assumed, Nature Geoscience, 10, 79, 10.1038/ngeo2882, 2017.

Arora, V. K., Boer, G. J., Friedlingstein, P., Eby, M., Jones, C. D., Christian, J. R., Bonan, G., Bopp, L., Brovkin, V., Cadule, P., Hajima, T., Ilyina, T., Lindsay, K., Tjiputra, J. F., and Wu, T.: Carbon-concentration and carbon-climate feedbacks in CMIP5 earth system models, Journal of Climate, 26, 5289-5314, 10.1175/JCLI-D-12-00494.1, 2013.

Arora, V. K., Katavouta, A., Williams, R. G., Jones, C. D., Brovkin, V., Friedlingstein, P., Schwinger, J., Bopp, L., Boucher, O., Cadule, P., Chamberlain, M. A., Christian, J. R., Delire, C., Fisher, R. A., Hajima, T., Ilyina, T., Joetzjer, E., Kawamiya, M., Koven, C. D., Krasting, J. P., Law, R. M., Lawrence, D. M., Lenton, A., Lindsay, K., Pongratz, J., Raddatz, T., Séférian, R., Tachiiri, K., Tjiputra, J. F., Wiltshire, A., Wu, T., and Ziehn, T.: Carbon–concentration and carbon–climate feedbacks in CMIP6 models and their comparison to CMIP5 models, Biogeosciences, 17, 4173-4222, 10.5194/bg-17-4173-2020, 2020.

Bastos, A., Ciais, P., Barichivich, J., Bopp, L., Brovkin, V., Gasser, T., Peng, S., Pongratz, J., Viovy, N., and Trudinger, C. M.: Re-evaluating the 1940s CO2 plateau, Biogeosciences, 13, 4877-4897, 10.5194/bg-13-4877-2016, 2016.

Bloch-Johnson, J., Pierrehumbert, R. T., and Abbot, D. S.: Feedback temperature dependence determines the risk of high warming, Geophysical Research Letters, 42, 4973-4980, https://doi.org/10.1002/2015GL064240, 2015.

Bonnet, R., Swingedouw, D., Gastineau, G., Boucher, O., Deshayes, J., Hourdin, F., Mignot, J., Servonnat, J., and Sima, A.: Increased risk of near term global warming due to a recent AMOC weakening, Nature Communications, 12, 6108, 10.1038/s41467-021-26370-0, 2021.

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 CO2 concentration changes, Environmental Research Letters, 7, 024013, 10.1088/1748-9326/7/2/024013, 2012.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S. K., Sherwood, S., Stevens, B., and Zhang, X. Y.: Clouds and Aerosols, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K.,

Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013a.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S. K., Sherwood, S., Stevens, B., and Zhang, X. Y.: Clouds and Aerosols, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 571-658, https://www.ipcc.ch/, 2013b.

Bright, R. M. and Kvalevåg, M. M.: Technical Note: Evaluating a simple parameterization of radiative shortwave forcing from surface albedo change, Atmos. Chem. Phys., 13, 11169-11174, 10.5194/acp-13-11169-2013, 2013.

Ciais, P., Gasser, T., Paris, J. D., Caldeira, K., Raupach, M. R., Canadell, J. G., Patwardhan, A., Friedlingstein, P., Piao, S. L., and Gitz, V.: Attributing the increase in atmospheric CO2 to emitters and absorbers, Nature Climate Change, 3, 926-930, 10.1038/nclimate1942, 2013a.

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., Jones, C., Le Quere, C., Myneni, R. B., Piao, S., and Thornton, P.: Carbon and Other Biogeochemical Cycles, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 465-570, https://www.ipcc.ch/, 2013b.

Clarke, L. E., Jiang, K., Akimoto, K., Babiker, M., Blanford, G., Fisher-Vanden, K., Hourcade, J.-C., Krey, V., Kriegler, E., Löschel, A., McCollum, D., Paltsev, S., Rose, S., Shukla, P. R., Tavoni, M., van der Zwaan, B. C. C., and van Vuuren, D. P.: Assessing transformation pathways, in: Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Edenhofer, O., Pichs-Madruga, R. S., Y., Farahani, E., Kadner, S., Seyboth, K., Adler, A., Baum, I., Brunner, S., Eickemeier, P., Kriemann, B., Savolainen, J., Schlömer, S., von Stechow, C., T., Z., and Minx, J. C., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 413-510, https://www.ipcc.ch/, 2014.

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W. J., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C., Weaver, A. J., and Wehner, M.: Long-term Climate Change: Projections, Commitments and Irreversibility, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1029-1136, https://www.ipcc.ch/, 2013.

Collins, W. J., Lamarque, J. F., Schulz, M., Boucher, O., Eyring, V., Hegglin, M. I., Maycock, A., Myhre, G., Prather, M., Shindell, D., and Smith, S. J.: AerChemMIP: quantifying the effects of chemistry and aerosols in CMIP6, Geosci. Model Dev., 10, 585-607, 10.5194/gmd-10-585-2017, 2017.

Cowtan, K. and Way, R. G.: Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends, Quarterly Journal of the Royal Meteorological Society, 140, 1935-1944, 10.1002/qj.2297, 2013.

Crichton, K. A., Roche, D. M., Krinner, G., and Chappellaz, J.: A simplified permafrost-carbon model for long-term climate studies with the CLIMBER-2 coupled earth system model, Geosci. Model Dev., 7, 3111-3134, 10.5194/gmd-7-3111-2014, 2014.

Dong, Y., Armour, K. C., Zelinka, M. D., Proistosescu, C., Battisti, D. S., Zhou, C., and Andrews, T.: Intermodel Spread in the Pattern Effect and Its Contribution to Climate Sensitivity in CMIP5 and CMIP6 Models, Journal of Climate, 33, 7755-7775, 10.1175/JCLI-D-19-1011.1, 2020.

Dorheim, K., Smith, S. J., and Bond-Lamberty, B.: HIRM v1.0: a hybrid impulse response model for climate modeling and uncertainty analyses, Geosci. Model Dev., 14, 365-375, 10.5194/gmd-14-365-2021, 2021.

Ehhalt, D., Prather, M., Dentener, M., Derwent, F., Dlugokencky, R., Holland, E., Isaksen, I., Katima, J., Kirchhoff, V., Matson, P., Midgley, P., and Wang, M.: Atmospheric Chemistry and Greenhouse Gases, in: Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Houghton, J. T., Ding, T., Griggs, D. J., Noguer, M., van der Linden, P. J., Dai, X., Maskell, K., and Johnson, C. A., Cambridge University Press, Cambridge, UK and New York, NY, USA, 239–287, https://www.ipcc.ch, 2001.

ESGF: ESGF, input4MIPs, https://esgf-node.llnl.gov/search/input4mips/, 2018.

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

Eyring, V., Gillett, N. P., Achuta Rao, K. M., Barimalala, R., Barreiro Parrillo, M., Bellouin, N., Cassou, C., Durack, P. J., Kosaka, Y., McGregor, S., Min, S., Morgenstern, O., and Sun, Y.: Human Influence on the Climate System, 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., 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, United Kingdom and New York, NY, USA, 423–552, https://www.ipcc.ch/, 2021.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., and Rummukainen, M.: Evaluation of Climate Models, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M.,

Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 741-866, https://www.ipcc.ch/, 2013.

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., von Bloh, W., Brovkin, V., Cadule, P., Doney, S., Eby, M., Fung, I., Bala, G., John, J., Jones, C., Joos, F., Kato, T., Kawamiya, M., Knorr, W., Lindsay, K., Matthews, H. D., Raddatz, T., Rayner, P., Reick, C., Roeckner, E., Schnitzler, K. G., Schnur, R., Strassmann, K., Weaver, A. J., Yoshikawa, C., and Zeng, N.: Climate–Carbon Cycle Feedback Analysis: Results from the C4MIP Model Intercomparison, Journal of Climate, 19, 3337-3353, 10.1175/JCLI3800.1, 2006.

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S., Aragão, L. E. O. C., Arneth, A., Arora, V., Bates, N. R., Becker, M., Benoit-Cattin, A., Bittig, H. C., Bopp, L., Bultan, S., Chandra, N., Chevallier, F., Chini, L. P., Evans, W., Florentie, L., Forster, P. M., Gasser, T., Gehlen, M., Gilfillan, D., Gkritzalis, T., Gregor, L., Gruber, N., Harris, I., Hartung, K., Haverd, V., Houghton, R. A., Ilyina, T., Jain, A. K., Joetzjer, E., Kadono, K., Kato, E., Kitidis, V., Korsbakken, J. I., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Liu, Z., Lombardozzi, D., Marland, G., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S. I., Niwa, Y., O'Brien, K., Ono, T., Palmer, P. I., Pierrot, D., Poulter, B., Resplandy, L., Robertson, E., Rödenbeck, C., Schwinger, J., Séférian, R., Skjelvan, I., Smith, A. J. P., Sutton, A. J., Tanhua, T., Tans, P. P., Tian, H., Tilbrook, B., van der Werf, G., Vuichard, N., Walker, A. P., Wanninkhof, R., Watson, A. J., Willis, D., Wiltshire, A. J., Yuan, W., Yue, X., and Zaehle, S.: Global Carbon Budget 2020, Earth Syst. Sci. Data, 12, 3269-3340, 10.5194/essd-12-3269-2020, 2020.

Friedrich, T., Timmermann, A., Tigchelaar, M., Timm, O. E., and Ganopolski, A.: Nonlinear climate sensitivity and its implications for future greenhouse warming, Science Advances, 2, e1501923, doi:10.1126/sciadv.1501923, 2016.

Fu, B., Gasser, T., Li, B., Tao, S., Ciais, P., Piao, S., Balkanski, Y., Li, W., Yin, T., Han, L., Li, X., Han, Y., An, J., Peng, S., and Xu, J.: Short-lived climate forcers have long-term climate impacts via the carbon–climate feedback, Nature Climate Change, 10, 851-855, 10.1038/s41558-020-0841-x, 2020.

Gasser, T.: Attribution régionalisée des causes anthropiques du changement climatique, Université Pierre et Marie Curie - Paris VI, 194 pp., http://www.theses.fr/2014PA066543, 2014.

Gasser, T. and Ciais, P.: A theoretical framework for the net land-to-atmosphere CO2 flux and its implications in the definition of "emissions from land-use change", Earth Syst. Dynam., 4, 171-186, 10.5194/esd-4-171-2013, 2013a.

Gasser, T. and Ciais, P.: A theoretical framework for the net land-to-atmosphere CO2 flux and its implications in the definition of "emissions from land-use change", Earth System Dynamics, 4, 171-186, 10.5194/esd-4-171-2013, 2013b.

Gasser, T., Guivarch, C., Tachiiri, K., Jones, C. D., and Ciais, P.: Negative emissions physically needed to keep global warming below 2 °C, Nat Commun, 6, 10.1038/ncomms8958, 2015.

Gasser, T., Crepin, L., Quilcaille, Y., Houghton, R. A., Ciais, P., and Obersteiner, M.: Historical CO2 emissions from land use and land cover change and their uncertainty, Biogeosciences, 17, 4075-4101, 10.5194/bg-17-4075-2020, 2020a.

Gasser, T., Crepin, L., Quilcaille, Y., Houghton, R. A., Ciais, P., and Obersteiner, M.: Historical CO2 emissions from land-use and land-cover change and their uncertainty, Biogeosciences Discuss., under review, 1-43, 10.5194/bg-2020-33, 2020b.

Gasser, T., Ciais, P., Boucher, O., Quilcaille, Y., Tortora, M., Bopp, L., and Hauglustaine, D.: The compact Earth system model OSCAR v2.2: Description and first results, Geoscientific Model Development, 10, 271-319, 10.5194/gmd-10-271-2017, 2017.

Gasser, T., Kechiar, M., Ciais, P., Burke, E. J., Kleinen, T., Zhu, D., Huang, Y., Ekici, A., and Obersteiner, M.: Path-dependent reductions in CO2 emission budgets caused by permafrost carbon release, Nature Geoscience, 11, 830-835, 10.1038/s41561-018-0227-0, 2018.

Geoffroy, O., Saint-Martin, D., Bellon, G., Voldoire, A., Olivié, D. J. L., and Tytéca, S.: Transient Climate Response in a Two-Layer Energy-Balance Model. Part II: Representation of the Efficacy of Deep-Ocean Heat Uptake and Validation for CMIP5 AOGCMs, Journal of Climate, 26, 1859-1876, 10.1175/JCLI-D-12-00196.1, 2013a.

Geoffroy, O., Saint-Martin, D., Olivié, D. J. L., Voldoire, A., Bellon, G., and Tytéca, S.: Transient Climate Response in a Two-Layer Energy-Balance Model. Part I: Analytical Solution and Parameter Calibration Using CMIP5 AOGCM Experiments, Journal of Climate, 26, 1841-1857, 10.1175/jcli-d-12-00195.1, 2013b.

Gidden, M. J., Riahi, K., Smith, S. J., Fujimori, S., Luderer, G., Kriegler, E., van Vuuren, D. P., van den Berg, M., Feng, L., Klein, D., Calvin, K., Doelman, J. C., Frank, S., Fricko, O., Harmsen, M., Hasegawa, T., Havlik, P., Hilaire, J., Hoesly, R., Horing, J., Popp, A., Stehfest, E., and Takahashi, K.: Global emissions pathways under different socioeconomic scenarios for use in CMIP6: a dataset of harmonized emissions trajectories through the end of the century, Geosci. Model Dev., 12, 1443-1475, 10.5194/gmd-12-1443-2019, 2019.

Gillett, N. P., Arora, V. K., Matthews, D., and Allen, M. R.: Constraining the ratio of global warming to cumulative CO2 emissions using CMIP5 simulations, Journal of Climate, 26, 6844-6858, 10.1175/JCLI-D-12-00476.1, 2013.

Gillett, N. P., Shiogama, H., Funke, B., Hegerl, G., Knutti, R., Matthes, K., Santer, B. D., Stone, D., and Tebaldi, C.: The Detection and Attribution Model Intercomparison Project (DAMIP v1.0) contribution to CMIP6, Geosci. Model Dev., 9, 3685-3697, 10.5194/gmd-9-3685-2016, 2016.

Gillett, N. P., Kirchmeier-Young, M., Ribes, A., Shiogama, H., Hegerl, G. C., Knutti, R., Gastineau, G., John, J. G., Li, L., Nazarenko, L., Rosenbloom, N., Seland, Ø., Wu, T., Yukimoto, S., and Ziehn, T.: Constraining human contributions to observed warming since the pre-industrial period, Nature Climate Change, 11, 207-212, 10.1038/s41558-020-00965-9, 2021.

Gregory, J. M., Andrews, T., Ceppi, P., Mauritsen, T., and Webb, M. J.: How accurately can the climate sensitivity to CO2 be estimated from historical climate change?, Climate Dynamics, 54, 129-157, 10.1007/s00382-019-04991-y, 2020.

Gregory, J. M., Andrews, T., Good, P., Mauritsen, T., and Forster, P. M.: Small global-mean cooling due to volcanic radiative forcing, Climate Dynamics, 1-13, 10.1007/s00382-016-3055-1, 2016.

Gregory, J. M., Ingram, W. J., Palmer, M. A., Jones, G. S., Stott, P. A., Thorpe, R. B., Lowe, J. A., Johns, T. C., and Williams, K. D.: A new method for diagnosing radiative forcing and climate sensitivity, Geophysical Research Letters, 31, 10.1029/2003GL018747, 2004.

Hansen, J., Sato, M., Ruedy, R., Nazarenko, L., Lacis, A., Schmidt, G. A., Russell, G., Aleinov, I., Bauer, M., Bauer, S., Bell, N., Cairns, B., Canuto, V., Chandler, M., Cheng, Y., Del Genio, A., Faluvegi, G., Fleming, E., Friend, A., Hall, T., Jackman, C., Kelley, M., Kiang, N., Koch, D., Lean, J., Lerner, J., Lo, K., Menon, S., Miller, R., Minnis, P., Novakov, T., Oinas, V., Perlwitz, J., Perlwitz, J., Rind, D., Romanou, A., Shindell, D., Stone, P., Sun, S., Tausnev, N., Thresher, D., Wielicki, B., Wong, T., Yao, M., and Zhang, S.: Efficacy of climate forcings, Journal of Geophysical Research: Atmospheres, 110, D18104, 10.1029/2005JD005776, 2005.

Hartin, C. A., Patel, P., Schwarber, A., Link, R. P., and Bond-Lamberty, B. P.: A simple object-oriented and open-source model for scientific and policy analyses of the global climate system - Hector v1.0, Geoscientific Model Development, 8, 939-955, 10.5194/gmd-8-939-2015, 2015.

Hartmann, D. L., Klein Tank, A. M. G., Rusticucci, M., Alexander, L. V., Brönnimann, S., Charabi, Y., Dentener, F. J., Dlugokencky, E. J., Easterling, D. R., Kaplan, A., Soden, B. J., Thorne, P. W., Wild, M., and Zhai, P. M.: Observations: Atmosphere and Surface, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 159-254, https://www.ipcc.ch/, 2013.

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M., Bond, T. C., Dawidowski, L., Kholod, N., Kurokawa, J. I., Li, M., Liu, L., Lu, Z., Moura, M. C. P., O'Rourke, P. R., and Zhang, Q.: Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), Geosci. Model Dev., 11, 369-408, 10.5194/gmd-11-369-2018, 2018.

Holmes, C. D., Prather, M. J., Søvde, O. A., and Myhre, G.: Future methane, hydroxyl, and their uncertainties: key climate and emission parameters for future predictions, Atmos. Chem. Phys., 13, 285-302, 10.5194/acp-13-285-2013, 2013.

Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., Doelman, J. C., Fisk, J., Fujimori, S., Klein Goldewijk, K., Hasegawa, T., Havlik, P., Heinimann, A., Humpenöder, F., Jungclaus, J., Kaplan, J. O., Kennedy, J., Krisztin, T., Lawrence, D., Lawrence, P., Ma, L., Mertz, O., Pongratz, J., Popp, A., Poulter, B., Riahi, K., Shevliakova, E., Stehfest, E., Thornton, P., Tubiello, F. N., van Vuuren, D. P., and Zhang, X.: Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6, Geosci. Model Dev., 13, 5425-5464, 10.5194/gmd-13-5425-2020, 2020.

IIASA: RCP database, http://tntcat.iiasa.ac.at:8787/RcpDb/dsd?Action=htmlpage&page=welcome, 2018a.

IIASA: SSP database, https://tntcat.iiasa.ac.at/SspDb/dsd?Action=htmlpage&page=welcome, 2018b.

IPCC: Annex II: Climate System Scenario Tables in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T.

F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1395–1446, 10.1017/CBO9781107415324.030, 2013.

IPCC: Global warming of 1.5C. An IPCC Special Report on the impacts of global warming of 1.5C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty, Cambridge University Press, Cambridge, UK and New York, NY, USA, 10.1017/9781009157940, 2018.

Jones, C., Robertson, E., Arora, V., Friedlingstein, P., Shevliakova, E., Bopp, L., Brovkin, V., Hajima, T., Kato, E., Kawamiya, M., Liddicoat, S., Lindsay, K., Reick, C. H., Roelandt, C., Segschneider, J., and Tjiputra, J.: Twenty-First-Century Compatible CO2 Emissions and Airborne Fraction Simulated by CMIP5 Earth System Models under Four Representative Concentration Pathways, Journal of Climate, 26, 4398-4413, 10.1175/jcli-d-12-00554.1, 2013.

Jones, C. D., Frölicher, T. L., Koven, C., MacDougall, A. H., Matthews, H. D., Zickfeld, K., Rogelj, J., Tokarska, K. B., Gillett, N. P., Ilyina, T., Meinshausen, M., Mengis, N., Séférian, R., Eby, M., and Burger, F. A.: The Zero Emissions Commitment Model Intercomparison Project (ZECMIP) contribution to C4MIP: quantifying committed climate changes following zero carbon emissions, Geosci. Model Dev., 12, 4375-4385, 10.5194/gmd-12-4375-2019, 2019.

Jones, C. D., Arora, V., Friedlingstein, P., Bopp, L., Brovkin, V., Dunne, J., Graven, H., Hoffman, F., Ilyina, T., John, J. G., Jung, M., Kawamiya, M., Koven, C., Pongratz, J., Raddatz, T., Randerson, J. T., and Zaehle, S.: C4MIP-The Coupled Climate-Carbon Cycle Model Intercomparison Project: Experimental protocol for CMIP6, Geoscientific Model Development, 9, 2853-2880, 10.5194/gmd-9-2853-2016, 2016.

Joos, F., Bruno, M., Fink, R., Siegenthaler, U., Stocker, T. F., Le Quéré, C., and Sarmiento, J. L.: An efficient and accurate representation of complex oceanic and biospheric models of anthropogenic carbon uptake, Tellus B: Chemical and Physical Meteorology, 48, 394-417, 10.3402/tellusb.v48i3.15921, 1996.

Keller, D. P., Lenton, A., Scott, V., Vaughan, N. E., Bauer, N., Ji, D., Jones, C. D., Kravitz, B., Muri, H., and Zickfeld, K.: The Carbon Dioxide Removal Model Intercomparison Project (CDR-MIP): Rationale and experimental design, Geoscientific Model Development Discussions, 1-72, 10.5194/gmd-11-1133-2018, 2018a.

Keller, D. P., Lenton, A., Scott, V., Vaughan, N. E., Bauer, N., Ji, D., Jones, C. D., Kravitz, B., Muri, H., and Zickfeld, K.: The Carbon Dioxide Removal Model Intercomparison Project (CDRMIP): rationale and experimental protocol for CMIP6, Geosci. Model Dev., 11, 1133-1160, 10.5194/gmd-11-1133-2018, 2018b.

Kravitz, B., Robock, A., Boucher, O., Schmidt, H., Taylor, K. E., Stenchikov, G., and Schulz, M.: The Geoengineering Model Intercomparison Project (GeoMIP), Atmospheric Science Letters, 12, 162-167, 10.1002/asl.316, 2011.

Kravitz, B., Robock, A., Tilmes, S., Boucher, O., English, J. M., Irvine, P. J., Jones, A., Lawrence, M. G., MacCracken, M., Muri, H., Moore, J. C., Niemeier, U., Phipps, S. J., Sillmann, J., Storelvmo, T., Wang, H., and Watanabe, S.: The Geoengineering Model Intercomparison Project Phase 6 (GeoMIP6): simulation design and preliminary results, Geosci. Model Dev., 8, 3379-3392, 10.5194/gmd-8-3379-2015, 2015.

Lawrence, D. M., Hurtt, G. C., Arneth, A., Brovkin, V., Calvin, K. V., Jones, A. D., Jones, C. D., Lawrence, P. J., Noblet-Ducoudre, N. D., Pongratz, J., Seneviratne, S. I., and Shevliakova, E.: The Land Use Model Intercomparison Project (LUMIP) contribution to CMIP6: Rationale and experimental design, Geoscientific Model Development, 9, 2973-2998, 10.5194/gmd-9-2973-2016, 2016.

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., Pickers, P. A., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Arneth, A., Arora, V. K., Barbero, L., Bastos, A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Doney, S. C., Gkritzalis, T., Goll, D. S., Harris, I., Haverd, V., Hoffman, F. M., Hoppema, M., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K., Johannessen, T., Jones, C. D., Kato, E., Keeling, R. F., Goldewijk, K. K., Landschützer, P., Lefèvre, N., Lienert, S., Liu, Z., Lombardozzi, D., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S., Neill, C., Olsen, A., Ono, T., Patra, P., Peregon, A., Peters, W., Peylin, P., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rocher, M., Rödenbeck, C., Schuster, U., Schwinger, J., Séférian, R., Skjelvan, I., Steinhoff, T., Sutton, A., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F. N., van der Laan-Luijkx, I. T., van der Werf, G. R., Viovy, N., Walker, A. P., Wiltshire, A. J., Wright, R., Zaehle, S., and Zheng, B.: Global Carbon Budget 2018, Earth Syst. Sci. Data, 10, 2141-2194, 10.5194/essd-10-2141-2018, 2018a.

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Pongratz, J., Manning, A. C., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Jackson, R. B., Boden, T. A., Tans, P. P., Andrews, O. D., Arora, V. K., Bakker, D. C. E., Barbero, L., Becker, M., Betts, R. A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Cosca, C. E., Cross, J., Currie, K., Gasser, T., Harris, I., Hauck, J., Haverd, V., Houghton, R. A., Hunt, C. W., Hurtt, G., Ilyina, T., Jain, A. K., Kato, E., Kautz, M., Keeling, R. F., Klein Goldewijk, K., Körtzinger, A., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Lima, I., Lombardozzi, D., Metzl, N., Millero, F., Monteiro, P. M. S., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S. I., Nojiri, Y., Padin, X. A., Peregon, A., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Reimer, J., Rödenbeck, C., Schwinger, J., Séférian, R., Skjelvan, I., Stocker, B. D., Tian, H., Tilbrook, B., Tubiello, F. N., van der Laan-Luijkx, I. T., van der Werf, G. R., van Heuven, S., Viovy, N., Vuichard, N., Walker, A. P., Watson, A. J., Wiltshire, A. J., Zaehle, S., and Zhu, D.: Global Carbon Budget 2017, Earth Syst. Sci. Data, 10, 405-448, 10.5194/essd-10-405-2018, 2018b.

Lee, Y. H., Lamarque, J. F., Flanner, M. G., Jiao, C., Shindell, D. T., Berntsen, T., Bisiaux, M. M., Cao, J., Collins, W. J., Curran, M., Edwards, R., Faluvegi, G., Ghan, S., Horowitz, L. W., McConnell, J. R., Ming, J., Myhre, G., Nagashima, T., Naik, V., Rumbold, S. T., Skeie, R. B., Sudo, K., Takemura, T., Thevenon, F., Xu, B., and Yoon, J. H.: Evaluation of preindustrial to present-day black carbon and its albedo forcing from Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP), Atmos. Chem. Phys., 13, 2607-2634, 10.5194/acp-13-2607-2013, 2013.

Li, B., Gasser, T., Ciais, P., Piao, S., Tao, S., Balkanski, Y., Hauglustaine, D., Boisier, J.-P., Chen, Z., Huang, M., Li, L. Z., Li, Y., Liu, H., Liu, J., Peng, S., Shen, Z., Sun, Z., Wang, R., Wang, T., Yin, G., Yin, Y., Zeng, H., Zeng, Z., and Zhou, F.: The contribution of China's emissions to global climate forcing, Nature, 531, 357-361, 10.1038/nature17165, 2016.

LUH2: Official LUH2 Future Scenario datasets for CMIP6, http://luh.umd.edu/data.shtml, 2018.

Lurton, T., Balkanski, Y., Bastrikov, V., Bekki, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Contoux, C., Cozic, A., Cugnet, D., Dufresne, J.-L., Éthé, C., Foujols, M.-A., Ghattas, J., Hauglustaine, D., Hu, R.-M., Kageyama, M., Khodri, M., Lebas, N., Levavasseur, G., Marchand, M., Ottlé, C., Peylin, P., Sima, A., Szopa, S., Thiéblemont, R., Vuichard, N., and Boucher, O.: Implementation of the CMIP6 Forcing Data in the IPSL-CM6A-LR Model, Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2019MS001940, 2020.

MacDougall, A. H., Frölicher, T. L., Jones, C. D., Rogelj, J., Matthews, H. D., Zickfeld, K., Arora, V. K., Barrett, N. J., Brovkin, V., Burger, F. A., Eby, M., Eliseev, A. V., Hajima, T., Holden, P. B., Jeltsch-Thömmes, A., Koven, C., Menviel, L., Michou, M., Mokhov, I. I., Oka, A., Schwinger, J., Séférian, R., Shaffer, G., Sokolov, A., Tachiiri, K., Tjiputra, J., Wiltshire, A., and Ziehn, T.: Is there warming in the pipeline? A multi-model analysis of the zero emission commitment from CO2, Biogeosciences Discuss., 2020, 1-45, 10.5194/bg-2019-492, 2020.

Marvel, K., Schmidt, G. A., Miller, R. L., and Nazarenko, L. S.: Implications for climate sensitivity from the response to individual forcings, Nature Climate Change, 6, 386-389, 10.1038/nclimate2888, 2016.

Matthes, K., Funke, B., Andersson, M. E., Barnard, L., Beer, J., Charbonneau, P., Clilverd, M. A., Dudok De Wit, T., Haberreiter, M., Hendry, A., Jackman, C. H., Kretzschmar, M., Kruschke, T., Kunze, M., Langematz, U., Marsh, D. R., Maycock, A. C., Misios, S., Rodger, C. J., Scaife, A. A., Seppälä, A., Shangguan, M., Sinnhuber, M., Tourpali, K., Usoskin, I., Van De Kamp, M., Verronen, P. T., and Versick, S.: Solar forcing for CMIP6 (v3.2), Geoscientific Model Development, 10, 2247-2302, 10.5194/gmd-10-2247-2017, 2017.

McNeall, D., Williams, J., Booth, B., Betts, R., Challenor, P., Wiltshire, A., and Sexton, D.: The impact of structural error on parameter constraint in a climate model, Earth Syst. Dynam., 7, 917-935, 10.5194/esd-7-917-2016, 2016.

Meehl, G. A., Senior, C. A., Eyring, V., Flato, G., Lamarque, J.-F., Stouffer, R. J., Taylor, K. E., and Schlund, M.: Context for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models, Science Advances, 6, eaba1981, 10.1126/sciadv.aba1981, 2020.

Meinshausen, M., Raper, S. C. B., and Wigley, T. M. L.: Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6 – Part 1: Model description and calibration, Atmos. Chem. Phys., 11, 1417-1456, 10.5194/acp-11-1417-2011, 2011a.

Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M. L. T., Lamarque, J.-F., Matsumoto, K., Montzka, S. A., Raper, S. C. B., Riahi, K., Thomson, A., Velders, G. J. M., and van Vuuren, D. P. P.: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300, Climatic Change, 109, 213-241, 10.1007/s10584-011-0156-z, 2011b.

Meinshausen, M., Nicholls, Z., Lewis, J., Gidden, M. J., Vogel, E., Freund, M., Beyerle, U., Gessner, C., Nauels, A., Bauer, N., Canadell, J. G., Daniel, J. S., John, A., Krummel, P., Luderer, G., Meinshausen, N., Montzka, S. A., Rayner, P., Reimann, S., Smith, S. J., van den Berg, M., Velders, G. J. M., Vollmer, M., and Wang, H. J.: The SSP greenhouse gas concentrations and their extensions to 2500, Geosci. Model Dev. Discuss., 2019, 1-77, 10.5194/gmd-2019-222, 2019.

Meinshausen, M., Vogel, E., Nauels, A., Lorbacher, K., Meinshausen, N., Etheridge, D. M., Fraser, P. J., Montzka, S. A., Rayner, P. J., Trudinger, C. M., Krummel, P. B., Beyerle, U., Canadell, J. G., Daniel, J. S., Enting, I. G., Law, R. M.,

Lunder, C. R., O'Doherty, S., Prinn, R. G., Reimann, S., Rubino, M., Velders, G. J. M., Vollmer, M. K., Wang, R. H. J., and Weiss, R.: Historical greenhouse gas concentrations for climate modelling (CMIP6), Geoscientific Model Development, 10, 2057-2116, 10.5194/gmd-10-2057-2017, 2017.

Melnikova, I., Boucher, O., Cadule, P., Tanaka, K., Gasser, T., Hajima, T., Quilcaille, Y., Shiogama, H., Séférian, R., Tachiiri, K., Vuichard, N., Yokohata, T., and Ciais, P.: Impact of bioenergy crops expansion on climate-carbon cycle feedbacks in overshoot scenarios, Earth Syst. Dynam. Discuss., 2021, 1-22, 10.5194/esd-2021-72, 2021.

Melton, J. R., Wania, R., Hodson, E. L., Poulter, B., Ringeval, B., Spahni, R., Bohn, T., Avis, C. A., Beerling, D. J., Chen, G., Eliseev, A. V., Denisov, S. N., Hopcroft, P. O., Lettenmaier, D. P., Riley, W. J., Singarayer, J. S., Subin, Z. M., Tian, H., Zürcher, S., Brovkin, V., van Bodegom, P. M., Kleinen, T., Yu, Z. C., and Kaplan, J. O.: Present state of global wetland extent and wetland methane modelling: conclusions from a model inter-comparison project (WETCHIMP), Biogeosciences, 10, 753-788, 10.5194/bg-10-753-2013, 2013.

Morice, C. P., Kennedy, J. J., Rayner, N. A., and Jones, P. D.: Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set, Journal of Geophysical Research: Atmospheres, 117, 10.1029/2011JD017187, 2012.

Myhre, G., Samset, B. H., Schulz, M., Balkanski, Y., Bauer, S., Berntsen, T. K., Bian, H., Bellouin, N., Chin, M., Diehl, T., Easter, R. C., Feichter, J., Ghan, S. J., Hauglustaine, D., Iversen, T., Kinne, S., Kirkevåg, A., Lamarque, J. F., Lin, G., Liu, X., Lund, M. T., Luo, G., Ma, X., van Noije, T., Penner, J. E., Rasch, P. J., Ruiz, A., Seland, Ø., Skeie, R. B., Stier, P., Takemura, T., Tsigaridis, K., Wang, P., Wang, Z., Xu, L., Yu, H., Yu, F., Yoon, J. H., Zhang, K., Zhang, H., and Zhou, C.: Radiative forcing of the direct aerosol effect from AeroCom Phase II simulations, Atmos. Chem. Phys., 13, 1853-1877, 10.5194/acp-13-1853-2013, 2013.

Newman, P. A., Daniel, J. S., Waugh, D. W., and Nash, E. R.: A new formulation of equivalent effective stratospheric chlorine (EESC), Atmos. Chem. Phys., 7, 4537-4552, 10.5194/acp-7-4537-2007, 2007.

Nicholls, Z., Meinshausen, M., Lewis, J., Corradi, M. R., Dorheim, K., Gasser, T., Gieseke, R., Hope, A. P., Leach, N. J., McBride, L. A., Quilcaille, Y., Rogelj, J., Salawitch, R. J., Samset, B. H., Sandstad, M., Shiklomanov, A., Skeie, R. B., Smith, C. J., Smith, S. J., Su, X., Tsutsui, J., Vega-Westhoff, B., and Woodard, D. L.: Reduced Complexity Model Intercomparison Project Phase 2: Synthesising Earth system knowledge for probabilistic climate projections, Earth's Future, n/a, https://doi.org/10.1029/2020EF001900, 2021.

Nicholls, Z. R. J., Meinshausen, M., Lewis, J., Gieseke, R., Dommenget, D., Dorheim, K., Fan, C. S., Fuglestvedt, J. S., Gasser, T., Golüke, U., Goodwin, P., Kriegler, E., Leach, N. J., Marchegiani, D., Quilcaille, Y., Samset, B. H., Sandstad, M., Shiklomanov, A. N., Skeie, R. B., Smith, C. J., Tanaka, K., Tsutsui, J., and Xie, Z.: Reduced complexity model intercomparison project phase 1: Protocol, results and initial observations, Geosci. Model Dev. Discuss., 2020, 1-33, 10.5194/gmd-2019-375, 2020.

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, 10.5194/esd-11-737-2020, 2020.

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 (ScenarioMIP) for CMIP6, Geoscientific Model Development, 9, 3461-3482, 10.5194/gmd-9-3461-2016, 2016.

Prather, M. J., Hsu, J., DeLuca, N. M., Jackman, C. H., Oman, L. D., Douglass, A. R., Fleming, E. L., Strahan, S. E., Steenrod, S. D., Søvde, O. A., Isaksen, I. S. A., Froidevaux, L., and Funke, B.: Measuring and modeling the lifetime of nitrous oxide including its variability, Journal of Geophysical Research: Atmospheres, 120, 5693-5705, 10.1002/2015JD023267, 2015.

Quilcaille, Y., Gasser, T., Ciais, P., Lecocq, F., Janssens-Maenhout, G., and Mohr, S.: Uncertainty in projected climate change arising from uncertain fossil-fuel emission factors, Environmental Research Letters, 13, 044017, 10.1088/1748-9326/aab304, 2018.

Raupach, M. R., Gloor, M., Sarmiento, J. L., Canadell, J. G., Frölicher, T. L., Gasser, T., Houghton, R. A., Le Quéré, C., and Trudinger, C. M.: The declining uptake rate of atmospheric CO2 by land and ocean sinks, Biogeosciences, 11, 3453-3475, 10.5194/bg-11-3453-2014, 2014.

Ravishankara, A. R., Daniel, J. S., and Portmann, R. W.: Nitrous Oxide (N2O): The Dominant Ozone-Depleting Substance Emitted in the 21st Century, Science, 326, 123-125, 10.1126/science.1176985, 2009.

Reddy, M. S. and Boucher, O.: Climate impact of black carbon emitted from energy consumption in the world's regions, Geophysical Research Letters, 34, 10.1029/2006GL028904, 2007.

Riahi, K., van Vuuren, D. P., Kriegler, E., Edmonds, J., O'Neill, B. C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuaresma, J. C., Kc, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., Ebi, K., Hasegawa, T., Havlik, P., Humpenöder, F., Da Silva, L. A., Smith, S., Stehfest, E., Bosetti, V., Eom, J., Gernaat, D., Masui, T., Rogelj, J., Strefler, J., Drouet, L., Krey, V., Luderer, G., Harmsen, M., Takahashi, K., Baumstark, L., Doelman, J. C., Kainuma, M., Klimont, Z., Marangoni, G., Lotze-Campen, H., Obersteiner, M., Tabeau, A., and Tavoni, M.: The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview, Global Environmental Change, 42, 153-168, 10.1016/j.gloenvcha.2016.05.009, 2017.

Ribes, A., Qasmi, S., and Gillett, N. P.: Making climate projections conditional on historical observations, Science Advances, 7, eabc0671, 10.1126/sciadv.abc0671, 2021.

Rohling, E. J., Rohling, E. J., Sluijs, A., Dijkstra, H. A., Köhler, P., van de Wal, R. S. W., von der Heydt, A. S., Beerling, D. J., Berger, A., Bijl, P. K., Crucifix, M., DeConto, R., Drijfhout, S. S., Fedorov, A., Foster, G. L., Ganopolski, A., Hansen, J., Hönisch, B., Hooghiemstra, H., Huber, M., Huybers, P., Knutti, R., Lea, D. W., Lourens, L. J., Lunt, D., Masson-Delmotte, V., Medina-Elizalde, M., Otto-Bliesner, B., Pagani, M., Pälike, H., Renssen, H., Royer, D. L., Siddall, M., Valdes,

P., Zachos, J. C., Zeebe, R. E., and Members, P. P.: Making sense of palaeoclimate sensitivity, Nature, 491, 683-691, 10.1038/nature11574, 2012.

Schlund, M., Lauer, A., Gentine, P., Sherwood, S. C., and Eyring, V.: Emergent constraints on equilibrium climate sensitivity in CMIP5: do they hold for CMIP6?, Earth Syst. Dynam., 11, 1233-1258, 10.5194/esd-11-1233-2020, 2020.

Sellers, P. J., Bounoua, L., Collatz, G. J., Randall, D. A., Dazlich, D. A., Los, S. O., Berry, J. A., Fung, I., Tucker, C. J., Field, C. B., and Jensen, T. G.: Comparison of Radiative and Physiological Effects of Doubled Atmospheric CO2 on Climate, Science, 271, 1402-1406, 10.1126/science.271.5254.1402, 1996.

Sherwood, S. C., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., Hegerl, G., Klein, S. A., Marvel, K. D., Rohling, E. J., Watanabe, M., Andrews, T., Braconnot, P., Bretherton, C. S., Foster, G. L., Hausfather, Z., von der Heydt, A. S., Knutti, R., Mauritsen, T., Norris, J. R., Proistosescu, C., Rugenstein, M., Schmidt, G. A., Tokarska, K. B., and Zelinka, M. D.: An Assessment of Earth's Climate Sensitivity Using Multiple Lines of Evidence, Reviews of Geophysics, 58, e2019RG000678, 10.1029/2019RG000678, 2020.

Shine, K. P., Allan, R. P., Collins, W. J., and Fuglestvedt, J. S.: Metrics for linking emissions of gases and aerosols to global precipitation changes, Earth Syst. Dynam., 6, 525-540, 10.5194/esd-6-525-2015, 2015.

Smith, C. J., Forster, P. M., Allen, M., Leach, N., Millar, R. J., Passerello, G. A., and Regayre, L. A.: FAIR v1.3: a simple emissions-based impulse response and carbon cycle model, Geosci. Model Dev., 11, 2273-2297, 10.5194/gmd-11-2273-2018, 2018.

Stevens, B.: Rethinking the Lower Bound on Aerosol Radiative Forcing, Journal of Climate, 28, 4794-4819, 10.1175/JCLI-D-14-00656.1, 2015.

Stevenson, D. S., Young, P. J., Naik, V., Lamarque, J. F., Shindell, D. T., Voulgarakis, A., Skeie, R. B., Dalsoren, S. B., Myhre, G., Berntsen, T. K., Folberth, G. A., Rumbold, S. T., Collins, W. J., MacKenzie, I. A., Doherty, R. M., Zeng, G., van Noije, T. P. C., Strunk, A., Bergmann, D., Cameron-Smith, P., Plummer, D. A., Strode, S. A., Horowitz, L., Lee, Y. H., Szopa, S., Sudo, K., Nagashima, T., Josse, B., Cionni, I., Righi, M., Eyring, V., Conley, A., Bowman, K. W., Wild, O., and Archibald, A.: Tropospheric ozone changes, radiative forcing and attribution to emissions in the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP), Atmos. Chem. Phys., 13, 3063-3085, 10.5194/acp-13-3063-2013, 2013.

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, Science Advances, 6, eaaz9549, 10.1126/sciadv.aaz9549, 2020.

Trudinger, C. and Enting, I.: Comparison of formalisms for attributing responsibility for climate change: Non-linearities in the Brazilian Proposal approach, Climatic Change, 68, 67-99, 10.1007/s10584-005-6012-2, 2005.

van den Hurk, B., Kim, H., Krinner, G., Seneviratne, S. I., Derksen, C., Oki, T., Douville, H., Colin, J., Ducharne, A., Cheruy, F., Viovy, N., Puma, M. J., Wada, Y., Li, W., Jia, B., Alessandri, A., Lawrence, D. M., Weedon, G. P., Ellis, R., Hagemann, S., Mao, J., Flanner, M. G., Zampieri, M., Materia, S., Law, R. M., and Sheffield, J.: LS3MIP (v1.0) contribution to CMIP6: the Land Surface, Snow and Soil moisture Model Intercomparison Project – aims, setup and expected outcome, Geosci. Model Dev., 9, 2809-2832, 10.5194/gmd-9-2809-2016, 2016.

van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S. J., and Rose, S. K.: The representative concentration pathways: an overview, Climatic Change, 109, 5, 10.1007/s10584-011-0148-z, 2011.

von Schuckmann, K., Cheng, L., Palmer, M. D., Hansen, J., Tassone, C., Aich, V., Adusumilli, S., Beltrami, H., Boyer, T., Cuesta-Valero, F. J., Desbruyères, D., Domingues, C., García-García, A., Gentine, P., Gilson, J., Gorfer, M., Haimberger, L., Ishii, M., Johnson, G. C., Killick, R., King, B. A., Kirchengast, G., Kolodziejczyk, N., Lyman, J., Marzeion, B., Mayer, M., Monier, M., Monselesan, D. P., Purkey, S., Roemmich, D., Schweiger, A., Seneviratne, S. I., Shepherd, A., Slater, D. A., Steiner, A. K., Straneo, F., Timmermans, M. L., and Wijffels, S. E.: Heat stored in the Earth system: where does the energy go?, Earth Syst. Sci. Data, 12, 2013-2041, 10.5194/essd-12-2013-2020, 2020.

Williamson, D. B. and Sansom, P. G.: How Are Emergent Constraints Quantifying Uncertainty and What Do They Leave Behind?, Bulletin of the American Meteorological Society, 100, 2571-2588, 10.1175/BAMS-D-19-0131.1, 2019.

Zanchettin, D., Khodri, M., Timmreck, C., Toohey, M., Schmidt, A., Gerber, E. P., Hegerl, G., Robock, A., Pausata, F. S. R., Ball, W. T., Bauer, S. E., Bekki, S., Dhomse, S. S., Le Grande, A. N., Mann, G. W., Marshall, L., Mills, M., Marchand, M., Niemeier, U., Poulain, V., Rozanov, E., Rubino, A., Stenke, A., Tsigaridis, K., and Tummon, F.: The Model Intercomparison Project on the climatic response to Volcanic forcing (VolMIP): Experimental design and forcing input data for CMIP6, Geoscientific Model Development, 9, 2701-2719, 10.5194/gmd-9-2701-2016, 2016.

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., Klein, S. A., and Taylor, K. E.: Causes of Higher Climate Sensitivity in CMIP6 Models, Geophysical Research Letters, 47, 10.1029/2019GL085782, 2020.

Zickfeld, K., Azevedo, D., Mathesius, S., and Matthews, H. D.: Asymmetry in the climate–carbon cycle response to positive and negative CO2 emissions, Nature Climate Change, 11, 613-617, 10.1038/s41558-021-01061-2, 2021.

Table 1. List of CMIP6 and RCMIP simulations run with OSCAR. Standard names are used, and full description of the experiments are provided in references. Every experiment that is a scenario has been run with its extension up to 2500. A spin-up of 1000 years is associated to each of the 8 control experiments.

MIP	Simulations
DECK (Eyring et al., 2016)	*1pctCO2, abrupt-4xCO2, esm-hist, historical, piControl, esm-piControl*
AerChemMIP (Collins et al., 2017)	*hist-1950HC, hist-piAer, hist-piNTCF, ssp370-lowNTCF*
C4MIP (Jones et al., 2016)	*1pctCO2-bgc, 1pctCO2-rad, esm-ssp585, hist-bgc, ssp534-over-bgc, ssp585-bgc*
CDRMIP (Keller et al., 2018a)	*1pctCO2-CDR, esm-pi-cdr-pulse, esm-pi-CO2pulse, esm-yr2010CO2-cdr-pulse, esm-yr2010CO2-CO2pulse, esm-yr2010CO2-control, esm-yr2010CO2-noemit, esm-ssp534-over, esm-ssp585-ssp126Lu, yr2010CO2*
DAMIP (Gillett et al., 2016)	*hist-aer, hist-CO2, hist-GHG, hist-nat, hist-sol, hist-stratO3, hist-volc, ssp245-aer, ssp245-CO2, ssp245-GHG, ssp245-nat, ssp245-sol, ssp245-stratO3, ssp245-volc*
LUMIP (Lawrence et al., 2016)	*esm-ssp585-ssp126Lu, hist-noLu, land-cClim, land-cCO2, land-crop-grass, land-hist, land-hist-altLu1, land-hist-altLu2, land-hist-altStartYear, land-noLu, land-noShiftCultivate, land-noWoodHarv, ssp126-ssp370Lu, ssp370-ssp126Lu, land-piControl, land-piControl-altLu1, land-piControl-altLu2, land-piControl-altStartYear*
GeoMIP (Kravitz et al., 2015)	*G1, G2, G6solar*
ScenarioMIP (O'Neill et al., 2016)	*ssp119, ssp126, ssp245, ssp370, ssp434, ssp460, ssp534-over, ssp585*
ZECMIP (Jones et al., 2019)	*esm-1pctCO2, esm-1pct-brch-750PgC, esm-1pct-brch-1000PgC, esm-1pct-brch-2000PgC, esm-bell-750PgC, esm-bell-1000PgC, esm-bell-2000PgC*
RCMIP (Nicholls et al., 2020)	*1pctCO2-4xext, abrupt-0p5xCO2, abrupt-2xCO2, esm-abrupt-4xCO2, esm-histcmip5, esm-rcp26, esm-rcp45, esm-rcp60, esm-rcp85, esm-ssp119, esm-ssp126, esm-ssp245, esm-ssp370, esm-ssp370-lowNTCF, esm-ssp434, esm-ssp460, historical-CMIP5, rcp26, rcp45, rcp60, rcp85, ssp585-ssp126Lu, esm-piControl-CMIP5, piControl-CMIP5*

Table 2: Metrics of the climate system (ECS, TCR and TCRE). Metrics are provided for OSCAR v3.1 constrained using observations, and unconstrained. Values are provided as mean ± standard deviation, median and the [5%-95%] confidence interval. As explained in section 3.2, the ECS in OSCAR may be calculated using its parameters, or simply as the temperature at the end of *abrupt-2xCO2*. These values are compared to the ECS of (Meehl et al., 2020). The same source provides the values for the TCR. The TCRE of CMIP5 is compared to (Gillett et al., 2013). Values from RCMIP phase 2 (Nicholls et al., 2021) come from different sources: (Sherwood et al., 2020) for the ECS, (Tokarska et al., 2020) for the TCR and (Arora et al., 2020) for the TCRE.

		OSCAR v3.1		CMIP5	CMIP6	RCMIP, phase 2
		Unconstrained	Constrained			
ECS (K)	Parameter value	3.17 ± 0.63 3.28 [2.36-4.25]	2.78 ± 0.47 2.63 [2.36-3.75]	3.2 ± 0.7	3.7 ± 1.1	3.10 [2.30-4.70]
	End of *abrupt-2xCO2*	2.74 ± 0.52 2.61 [2.02-3.67]	2.52 ± 0.33 2.45 [2.08-3.22]			
TCR (K)		1.78 ± 0.28 1.77 [1.37-2.26]	1.66 ± 0.16 1.62 [1.41-1.96]	1.8 ± 0.40	2.0 ± 0.4	1.64 [0.98-2.29]
TCRE (K 1000 PgC^{-1})		1.67 ± 0.40 1.63 [1.08-2.37]	1.44 ± 0.20 1.41 [1.15-1.82]	1.63 ± 0.48 [0.8-2.4]	1.77 ± 0.37	1.77 [1.03-2.51]

Table 3: Metrics of the carbon-cycle (β and *γ*) from the C4MIP experiments. Metrics are provided for OSCAR v3.1 constrained using observations, and unconstrained. As explained by (Arora et al., 2013), different values for the metrics are calculated depending on the combination of experiments used: R stands for radiative (*1pctCO2-rad*), B for biogeochemical (*1pctCO2-bgc*) and F for full (*1pctCO2*). The change in the land carbon stocks includes permafrost carbon. Results from CMIP5 and CMIP6 are provided by C4MIP (Arora et al., 2020).

Time	Model	Method β	β (PgC ppm^{-1})		Method γ	γ (PgC K^{-1})	
			Land	Ocean		Land	Ocean
2xCO2	OSCAR v3.1 constrained	R-B, B-F	1.26 ± 0.47	1.05 ± 0.03	R-B, R-F	-34.7 ± 18.9	-13.0 ± 0.7
		R-F	1.21 ± 0.44	1.00 ± 0.06	B-F	-43.2 ± 23.8	-21.6 ± 6.3
	OSCAR v3.1 unconstrained	R-B, B-F	1.14 ± 0.64	1.05 ± 0.03	R-B, R-F	-30.8 ± 20.5	-13.0 ± 0.7
		R-F	1.10 ± 0.61	1.00 ± 0.05	B-F	-37.6 ± 26.4	-21.0 ± 5.7
	CMIP5	B-F	1.15 ± 0.63	0.95 ± 0.07	B-F	-37.0 ± 25.5	-9.4 ± 2.7
	CMIP6	B-F	1.22 ± 0.40	0.91 ± 0.09	B-F	-34.1 ± 38.4	-8.6 ± 2.9
4xCO2	OSCAR v3.1 constrained	R-B, B-F	1.06 ± 0.41	0.94 ± 0.03	R-B, R-F	-47.7 ± 23.8	-17.7 ± 1.3
		R-F	0.95 ± 0.37	0.86 ± 0.08	B-F	-72.3 ± 37.4	-37.1 ± 13.6
	OSCAR v3.1 unconstrained	R-B, B-F	0.96 ± 0.57	0.94 ± 0.03	R-B, R-F	-43.3 ± 25.5	-17.7 ± 1.3
		R-F	0.87 ± 0.50	0.86 ± 0.07	B-F	-63.1 ± 41.5	-35.5 ± 12.4
	CMIP5	B-F	0.93 ± 0.49	0.82 ± 0.07	B-F	-57.9 ± 38.2	-17.3 ± 3.8
	CMIP6	B-F	0.97 ± 0.40	0.78 ± 0.07	B-F	-45.1 ± 50.6	-17.2 ± 4.9

Table 4: Attribution of historical and future climate change. These contributions come either from experiments in which only the concerned forcing was prescribed (DAMIP), or from experiments in which it was removed (other MIPs). In either cases, non-linearities are ignored.

	Experiments		GSAT w.r.t. 1850-1900 (K)		RF ($W.m^{-2}$)	
	2006-2015	2091-2100	2006-2015	2091-2100	2006-2015	2091-2100
All forcings	*historical*	*ssp245*	0.98 ± 0.19	2.53 ± 0.25	2.07 ± 0.42	4.62 ± 0.29
WMGHGs†	*hist-GHG*	*ssp245-GHG*	1.24 ± 0.12	2.67 ± 0.29	2.53 ± 0.13	4.73 ± 0.27
NTCFs‡	*hist-aer*	*ssp245-aer*	-0.26 ± 0.22	-0.15 ±0.12	-0.48 ± 0.36	-0.16 ±–0.12
id.	*historical - hist-piNTCF*	--	-0.25 ± 0.21	--	-0.46 ± 0.35	--
Natural forcings	*hist-nat*	*ssp245-nat*	~ 0.03	~ 0.01	~ 0.09	~ 0.00
CO_2	*hist-CO2*	*ssp245-CO2*	0.74 ± 0.07	2.03 ± 0.22	1.52 ± 0.09	3.70 ± 0.24
CO_2 radiative effect only	*historical - hist-bgc*	--	0.75 ± 0.08	--	1.55 ± 0.04	--
CFCs and HCFCs†	*historical - hist-1950HC*	--	0.13 ± 0.02	--	0.27 ± 0.03	--
Stratospheric O_3	*hist-stratO3*	*ssp245-stratO3*	-0.03 ± 0.03	-0.02 ±0.03	-0.07 ± 0.06	-0.02 ± 0.05–
Aerosols	*historical - hist-piAer*	--	-0.33 ± 0.20	--	-0.63 ± 0.33	--
Solar activity	*hist-sol*	*ssp245-sol*	~ 0.02	~ 0.01	~ 0.03	~ 0.02
Volcanic activity	*hist-volc*	*ssp245-volc*	~ 0.01	~ -0.01	~ 0.06	~ -0.02
Land–se change	*historical - hist-noLu*	--	-0.03 ± 0.03	--	-0.05 ± 0.05	--

† In these experiments, because the atmospheric concentration of WMGHGs is prescribed, the indirect effects on tropospheric O_3 (from CH_4), stratospheric H_2O (from CH_4) and stratospheric O_3 (from N_2O and halogenated compounds) are also included.

‡ The effects listed in the previous note on WMGHGs are excluded from this experiment. Tropospheric O_3 does vary~~, however~~, but only because of the emission of ozone precursors and not because of varying atmospheric CH_4. Black carbon deposition on snow is also included in this experiment.

Table 5: Projected atmospheric CO_2, RF and GSAT in SSPs. Concentration- and emission-driven experiments are shown and compared to available CMIP6 projections. Values in bold are assumptions or inputs. Experiments whose name start with *esm-* are emission-driven; others are concentration-driven. GSAT from CMIP6 are provided as mean and standard deviation as well, with the number of models available in parenthesis. Here, projections from OSCAR are constrained to observations, while CMIP6 results are raw, without any constraints (Tokarska et al., 2020).

experiments	models	ERF (W m^{-2})	GSAT w.r.t. 1850-1900 (K)				CO_2 (ppm)	
		2100	2041-2050	2091-2100	2291-2300	2491-2500	2100	2300
esm-ssp585	OSCAR	8.40 ± 0.57	2.02 ± 0.22	3.99 ± 0.40	6.31 ± 0.83	6.29 ± 0.88	1058 ±63	1729 ± 148
esm-ssp585	CMIP6		2.41 ± 1.67 (3)	5.14 ± 3.92 (2)				
ssp585	OSCAR	8.76 ± 0.50	2.04 ± 0.19	4.16 ± 0.38	7.05 ± 0.87	7.24 ± 0.93	1135	2162
ssp585	CMIP6		2.72 ± 1.51 (17)	6.19 ± 3.13 (17)	13.51 ± 5.87 (2)		**1135**	**2162**
esm-ssp370	OSCAR	7.04 ± 0.66	1.85 ± 0.25	3.32 ± 0.35	5.54 ± 0.74	5.56 ± 0.80	809 ± 47	1200 ± 109
ssp370	OSCAR	7.41 ± 0.58	1.87 ± 0.21	3.50 ± 0.32	6.24 ± 0.75	6.41 ± 0.81	867	1483
ssp370	CMIP6		2.51 ± 1.48 (18)	5.1 ± 2.84 (16)			**867**	**1483**
esm-ssp460	OSCAR	5.32 ± 0.50	1.80 ± 0.23	2.68 ± 0.30	3.43 ± 0.51	3.34 ± 0.55	629 ± 35	667 ± 49
ssp460	OSCAR	5.64 ± 0.40	1.82 ± 0.19	2.84 ± 0.27	3.91 ± 0.47	3.89 ± 0.50	668	769
ssp460	CMIP6		2.46 ± 1.28 (4)	4.24 ± 1.80 (4)			**668**	**769**
esm-ssp245	OSCAR	4.63 ± 0.43	1.72 ± 0.21	2.38 ± 0.28	2.59 ± 0.41	2.40 ± 0.42	578 ± 31	565 ± 35
ssp245	OSCAR	4.86 ± 0.31	1.75 ± 0.17	2.50 ± 0.25	2.92 ± 0.37	2.79 ± 0.37	603	621
ssp245	CMIP6		2.41 ± 1.33 (15)	3.63 ± 1.82 (15)			**603**	**621**
esm-ssp534-over	OSCAR	2.93 ± 0.37	2.00 ± 0.22	1.73 ± 0.25	1.16 ± 0.23	1.02 ± 0.23	458 ± 23	374 ± 12
ssp534-over	OSCAR	3.36 ± 0.27	2.04 ± 0.19	1.95 ± 0.22	1.40 ± 0.20	1.29 ± 0.19	497	398
ssp534-over	CMIP6		2.88 ± 0.84 (6)	3.08 ± 1.06 (6)	1.85 ± 0.66 (2)		**497**	**398**
esm-ssp434	OSCAR	3.45 ± 0.40	1.64 ± 0.20	1.87 ± 0.24	1.51 ± 0.28	1.44 ± 0.29	451 ± 21	371 ± 15
ssp434	OSCAR	3.70 ± 0.31	1.65 ± 0.17	2.00 ± 0.21	1.73 ± 0.24	1.68 ± 0.25	473	392
ssp434	CMIP6		2.36 ± 1.1 (5)	3.23 ± 1.32 (5)			**473**	**392**
esm-ssp126	OSCAR	2.66 ± 0.29	1.54 ± 0.18	1.49 ± 0.21	1.17 ± 0.20	1.02 ± 0.20	439 ± 18	381 ± 11
ssp126	OSCAR	2.80 ± 0.20	1.58 ± 0.15	1.58 ± 0.17	1.31 ± 0.18	1.21 ± 0.18	446	396
ssp126	CMIP6		2.21 ± 1.1 (17)	2.38 ± 1.17 (17)	1.68 ± 0.7 (2)		**446**	**396**
esm-ssp119	OSCAR	2.0 ± 0.25	1.39 ± 0.17	1.15 ± 0.17	0.71 ± 0.15	0.61 ± 0.15	383 ± 12	334 ± 6
ssp119	OSCAR	2.14 ± 0.18	1.44 ± 0.14	1.24 ± 0.15	0.82 ± 0.13	0.74 ± 0.13	394	342
ssp119	CMIP6		2.36 ± 1.07 (6)	2.12 ± 0.92 (2)			**394**	**342**

Table 6: Zero Emissions Commitments at 25, 50, 90 and 500 years after emission cease. Only the ZECs for the experiment *esm-1pct-brch-1000PgC* are shown here, for comparison to results of ZECMIP. The full evolution of this experiment is shown in Figure ~~Figure 8~~.

	ZEC_{25} (K)	ZEC_{50} (K)	ZEC_{90} (K)	ZEC_{500} (K)
OSCAR v3.1	-0.01 ± 0.07	-0.02 ± 0.09	-0.01 ± 0.11	-0.21 ± 0.13
ZECMIP (MacDougall et al., 2020)	-0.01 ± 0.15	-0.06 ± 0.19	-0.11 ± 0.23	

Formatted: Font: Not Bold

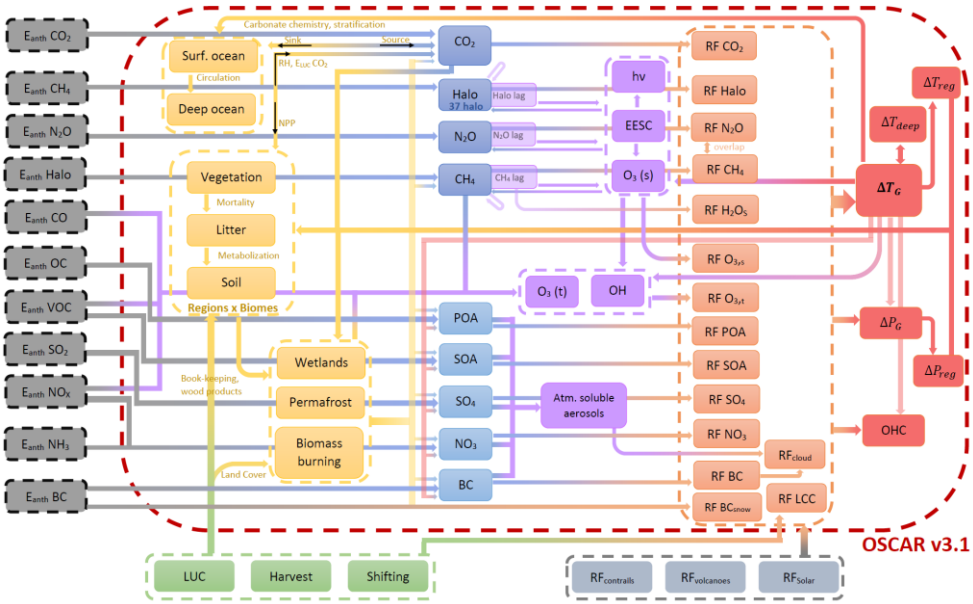

Figure 1: Conceptual figure of OSCAR v3.1. The central box with red dashed lines illustrates the framework of OSCAR v3.1, taking as inputs anthropogenic emissions (dark grey boxes), land use and land cover change (green boxes) and additional radiative forcings (light grey boxes). The components of OSCAR v3.1 are organized in this figure by category: ocean carbon, land carbon and other land processes are in yellow boxes, while atmospheric concentrations are in blue boxes, atmospheric chemistry in purple boxes, radiative forcings in orange boxes and climate system in red boxes. The complete description of OSCAR v2.2 is in (Gasser et al., 2017), while the update to OSCAR v3.1 is described in (Gasser et al., 2020a).

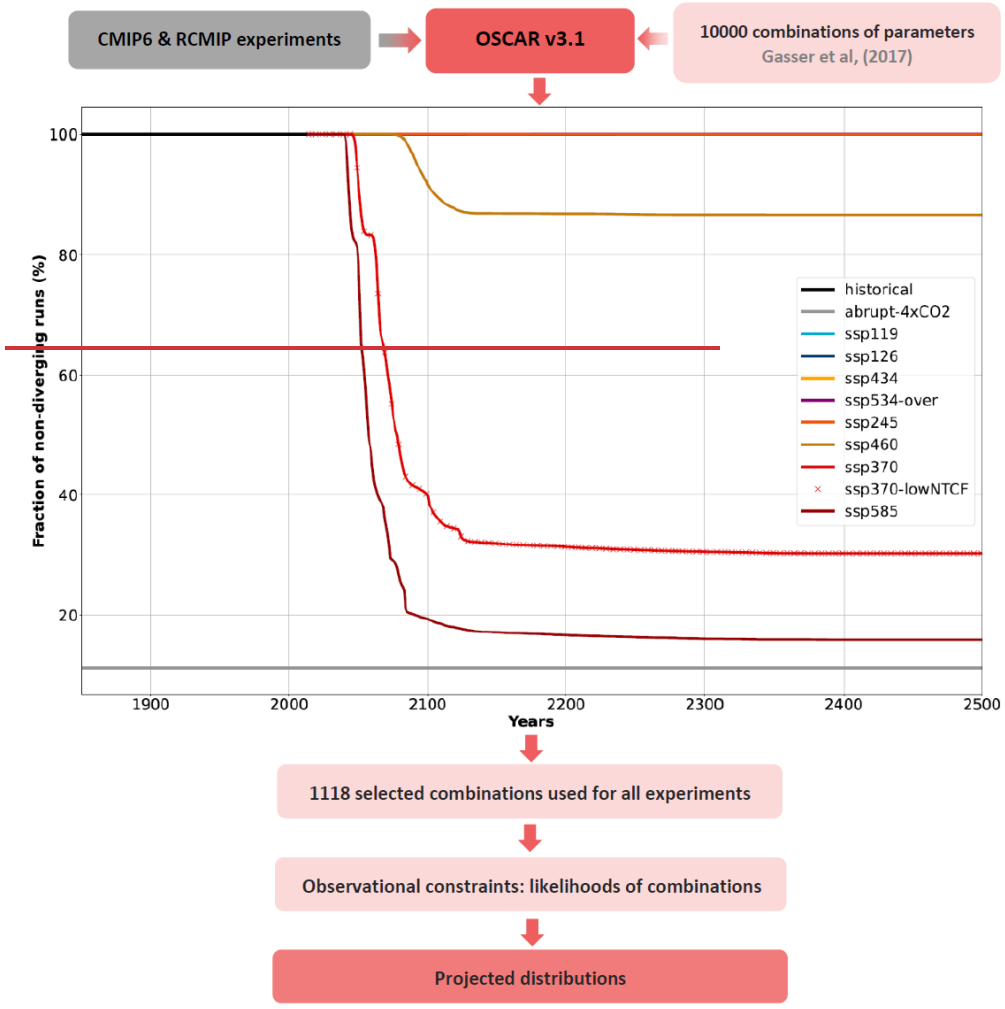

Figure 2: Conceptual description of the framework used in this study. The 10000 drawn configurations (Gasser et al., 2017) are used in OSCAR in a Monte-Carlo setup for all experiments. The exclusions are based on their exceedance to thresholds in the ocean sink, land sink, CO_2 emissions form LUC and CO_2 emissions from permafrost. The remaining subset common to each experiment is then used for all. The likelihood of the kept configurations is then calculated (Gasser et al., 2020a) and applied to all experiments.

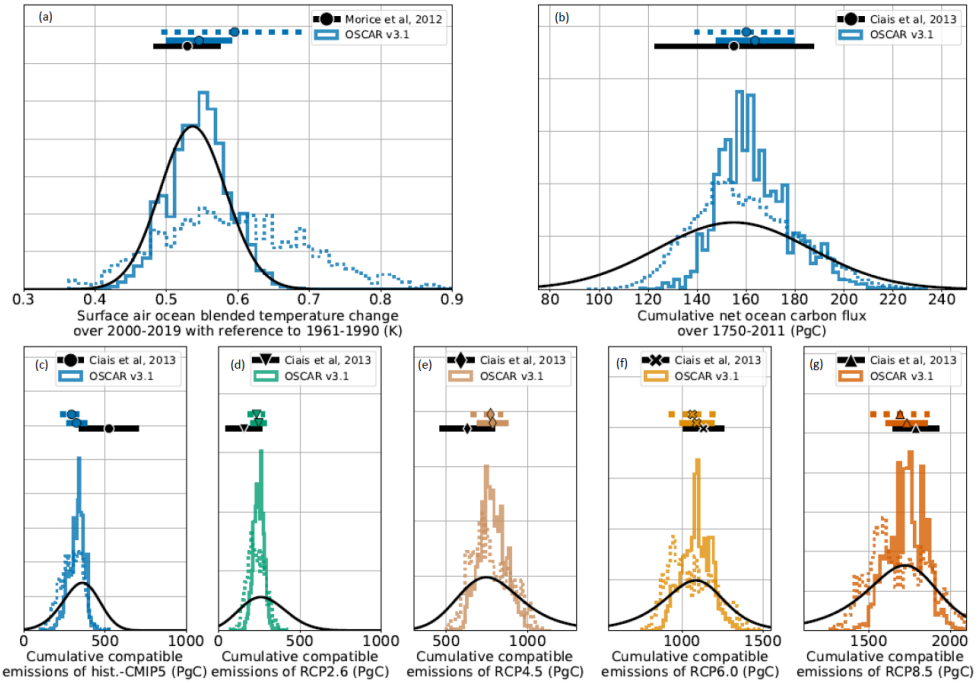

Figure ~~3~~2: Effect of the constraining step. The histograms are the results of OSCAR v3.1, with plain lines being for the constrained version, while the dotted lines are for the unconstrained version. Horizontal lines correspond to the average plus or minus one standard deviation. Cumulative compatible carbon emissions in PgC from *historical-CMIP5* are calculated over 1850-2011, while those of the RCPs are calculated over 2012-2100.

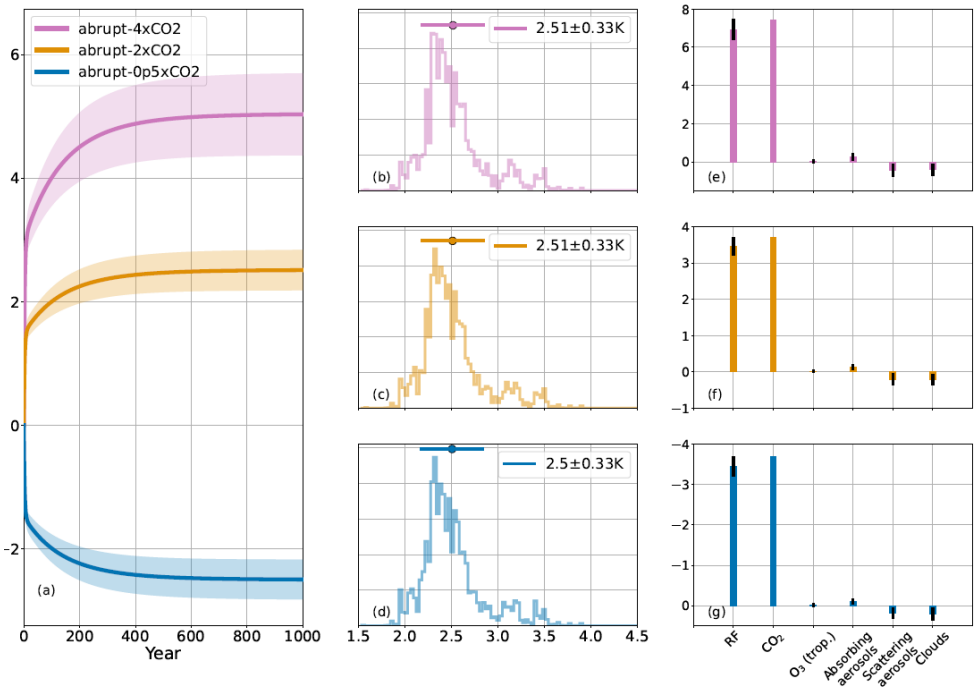

Figure 43: Abrupt idealized experiments. In the left panel, the plain lines represent the average change in surface air temperature, and its ± 1 standard deviation ranges using shaded areas. The three middle panels show the contributions to the total RF at equilibrium. Individual contributions from stratospheric O_3 and deposition of BC on snow are inferior to 0.1 $W.m^{-2}$ in the *abrupt-4xCO2* and have not been represented for clarity. The three right panels are the distributions of the ECS, calculated using equilibrium temperature, and thus including all the feedbacks of OSCAR. The horizontal plain line is the ECS average and ± 1 standard deviation range. These values with Pearson's moment coefficient of skewness are provided in the legend.

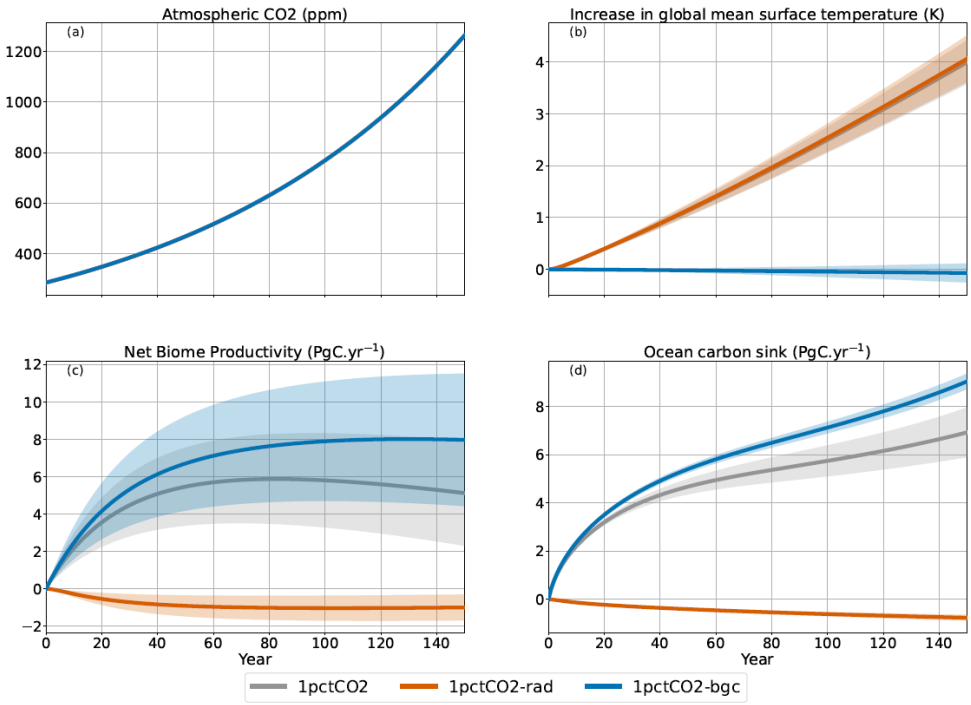

Figure 54: Experiments with 1% increase in the atmospheric CO_2. The plain lines are the averages, and the shaded areas represent ± 1 standard deviation ranges.

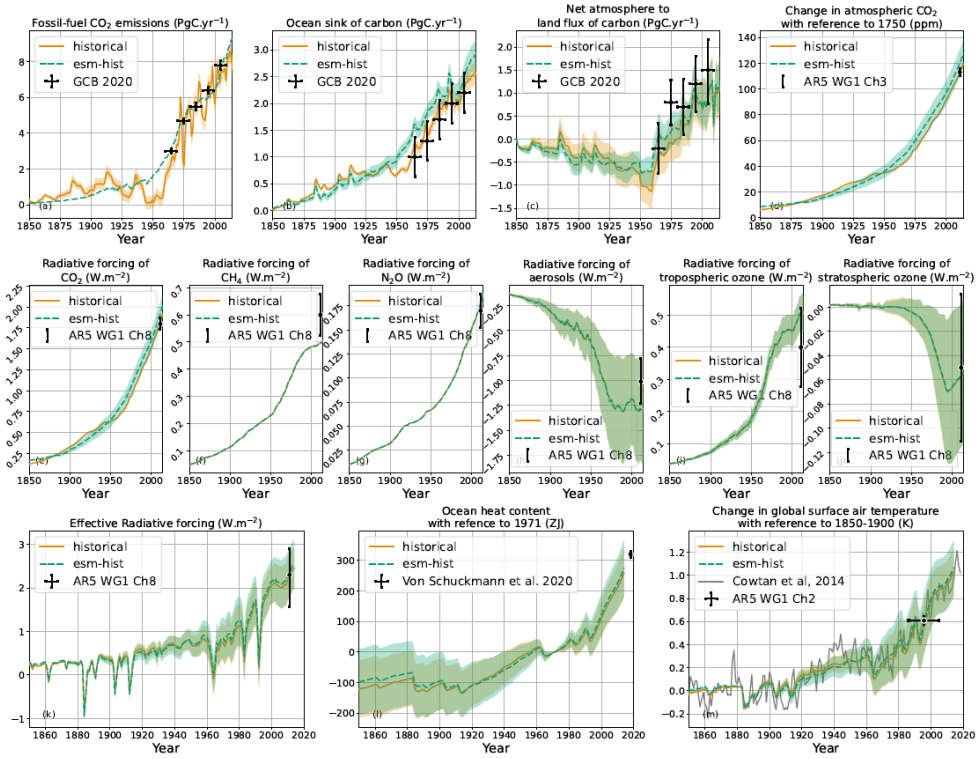

Figure 65: Emission- and concentration-driven historical scenarios. The plain lines are the averages, and the shaded areas represent ±–1 standard deviation ranges. The fossil-fuel CO2 emissions for the concentrations-driven historical are the compatible emissions, whereas those for the emissions-driven esm-hist are directly prescribed to OSCAR. Radiative forcings under esm-hist are not represented, for they are too close from the concentrations-driven historical. Radiative forcings are with respect to 1750. The sources for the observations are (Friedlingstein et al., 2020) for GCB2020, (Hartmann et al., 2013) for the 'AR5 WG1 Ch2', (Ciais et al., 2013b) for 'AR5 WG1 Ch3' and (Myhre et al., 2013) for 'AR5 WG1 Ch8'. The 90% ranges provided by AR5 are converted to ± 1 standard deviation ranges.

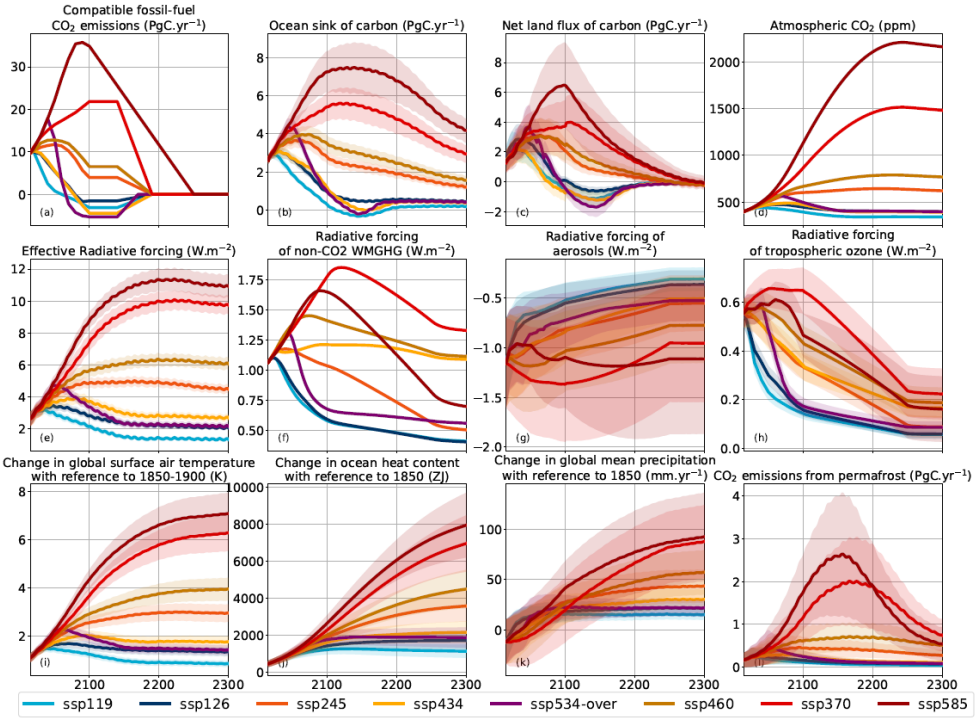

Figure ~~7~~6: Global projections following the main CMIP6 scenarios in concentration-driven mode. Extensions are shown only up to 2300. The lines are the averages, and the shaded areas represent ± 1 standard deviation ranges.

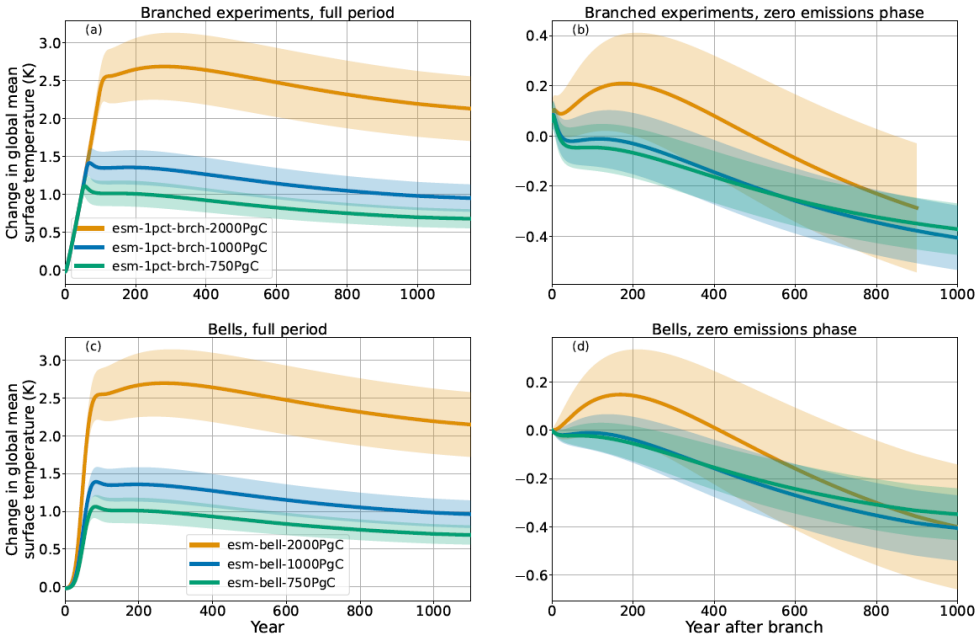

Figure 87: Change in global mean surface temperature for branched experiments (top panels) and bells experiments (bottom panels). The results over the zero-emission phase are shifted along the time axis so that $t = 0$ corresponds to the time of cessation of emission. The lines are the averages, and the shaded areas represent ± 1 standard deviation ranges.

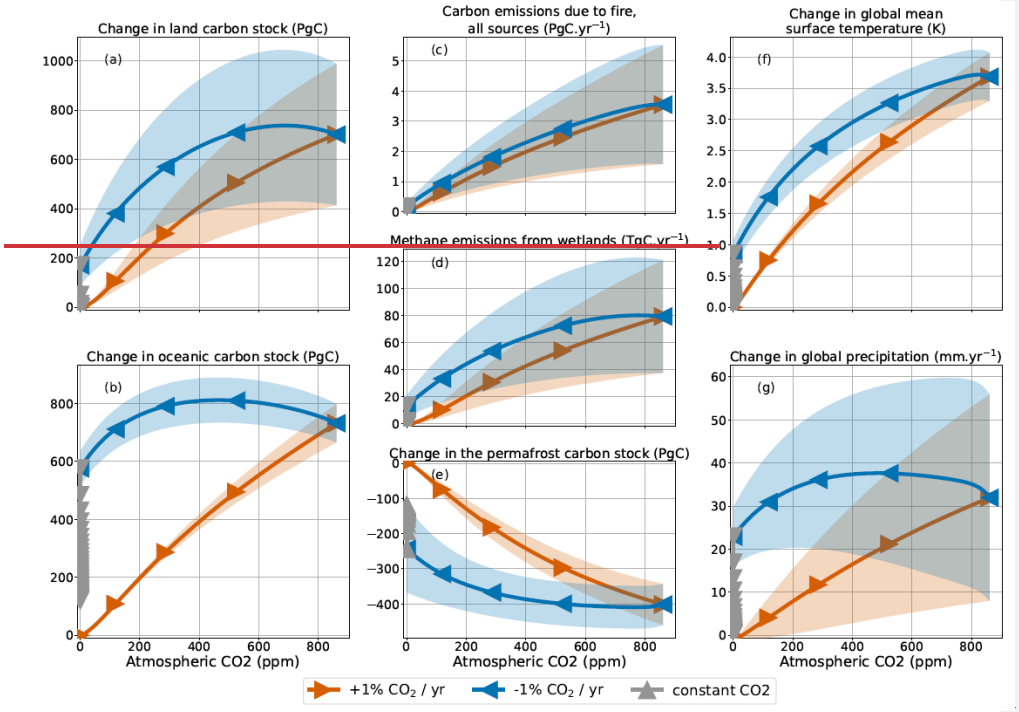

Figure A. 1: Reversibility experiment from CDRMIP. The orange lines correspond to the ramp-up of *1pctCO2-cdr*, the blue line to its ramp-down and the grey line to the 1000 years with constant atmospheric CO_2. The plain lines are the averages, and the shaded areas represent ± 1 standard deviation ranges.

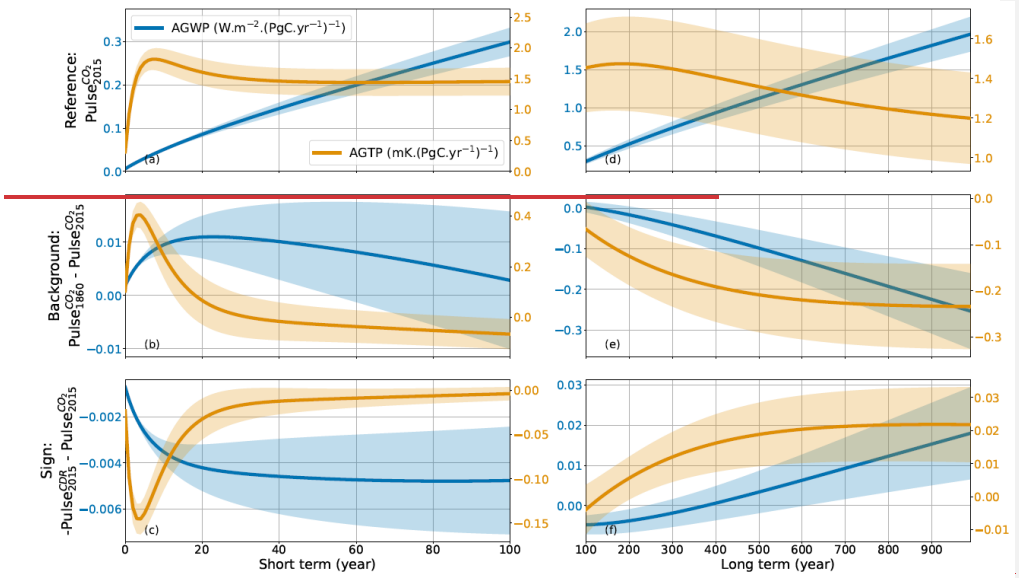

Figure A. 2: AGWP (blue) and AGTP (orange) of CO_2 for 100PgC of CO_2 emissions under actual environmental conditions. The dependency of this reference to a change of background is on the second line. The dependency to the sign of the pulse, emissions or removal, is on the third line. The lines are the averages, and the shaded areas represent ± 1 standard deviation ranges.

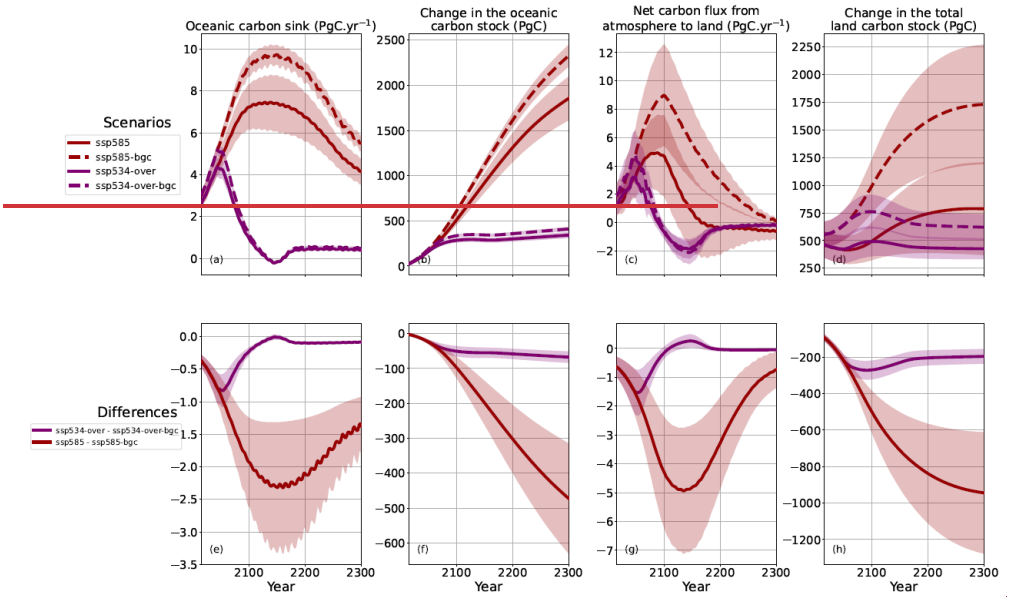

Figure A. 3: Effect of climate change on the carbon cycle in the scenarios *ssp534-over* and *ssp585*. The net flux from atmosphere from land is the sum of the land carbon sink, CO_2 emissions from land-use and land-cover change, and CO_2 and CH_4 emissions from permafrost. The changes in the total land carbon stock include those in the permafrost. Note that the increased uncertainty in the ocean sink before 2250 is an artefact of our exclusion procedure (see text on post-processing) that cannot capture the Monte Carlo members that already started diverging. Extensions are shown only up to 2300. The lines are the averages, and the shaded areas represent ± 1 standard deviation ranges.

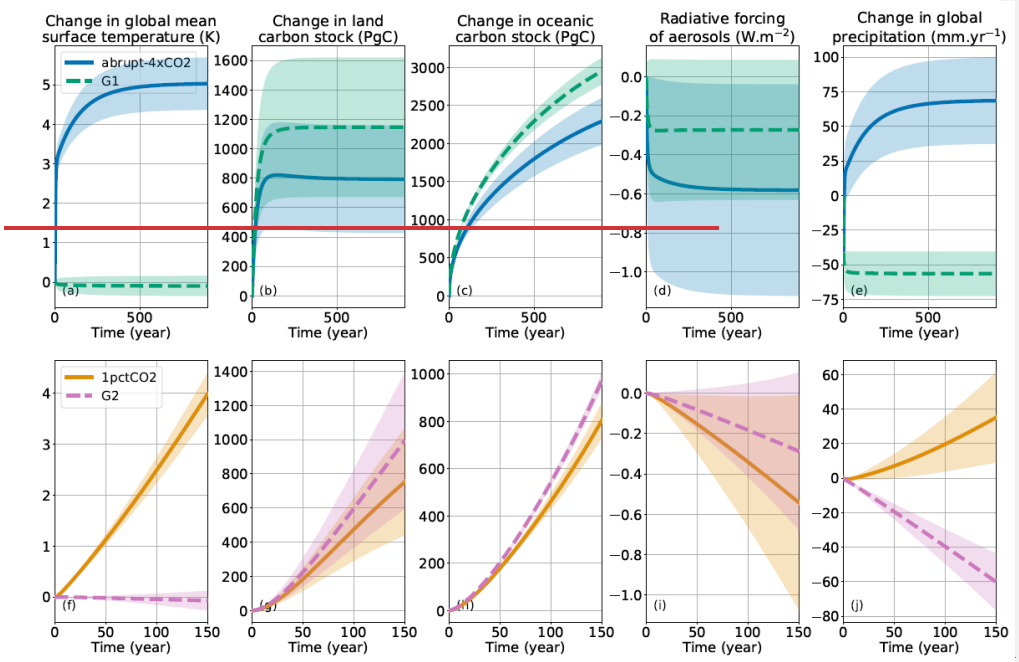

Figure A. 4: Experiments from GeoMIP compared to their DECK counterpart. The plain lines are the averages, and the shaded areas represent ± 1 standard deviation ranges.

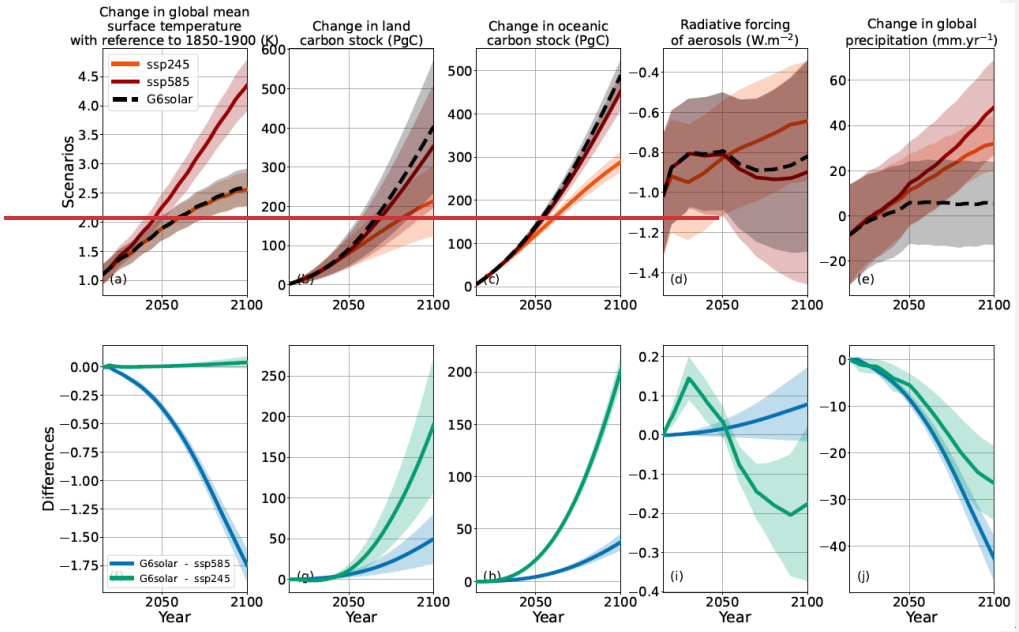

Figure A. 5: Effect of introducing SRM in the SSP5-8.5 to reach the SSP2-4.5. The lines are the averages, and the shaded areas represent ± 1 standard deviation ranges.

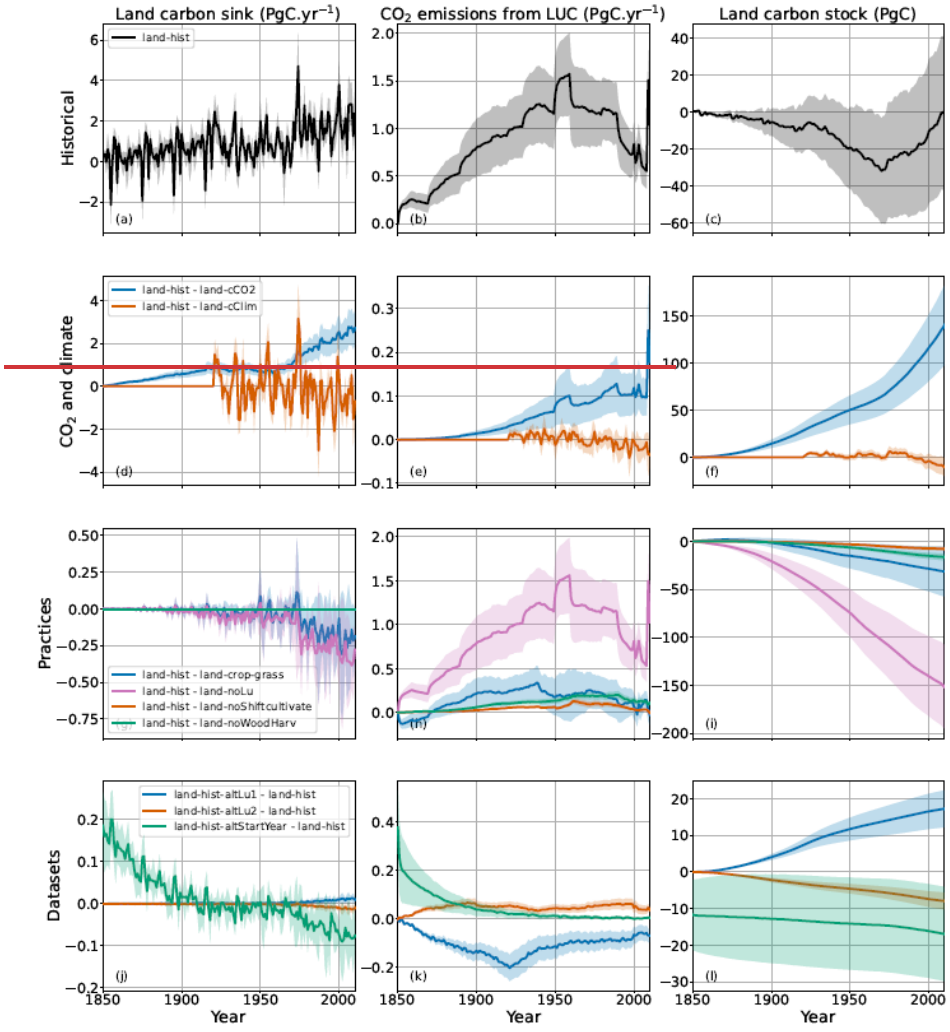

Figure A. 6: Land-use experiments from LUMIP. The first row of the figure corresponds to the reference experiment (*land-hist*) while other rows show sensitivity experiments as a difference to *land-hist*. *land-hist-altStartYear* is shown only from 1850 despite starting in 1700. The lines are the averages, and the shaded areas represent ± 1 standard deviation ranges.

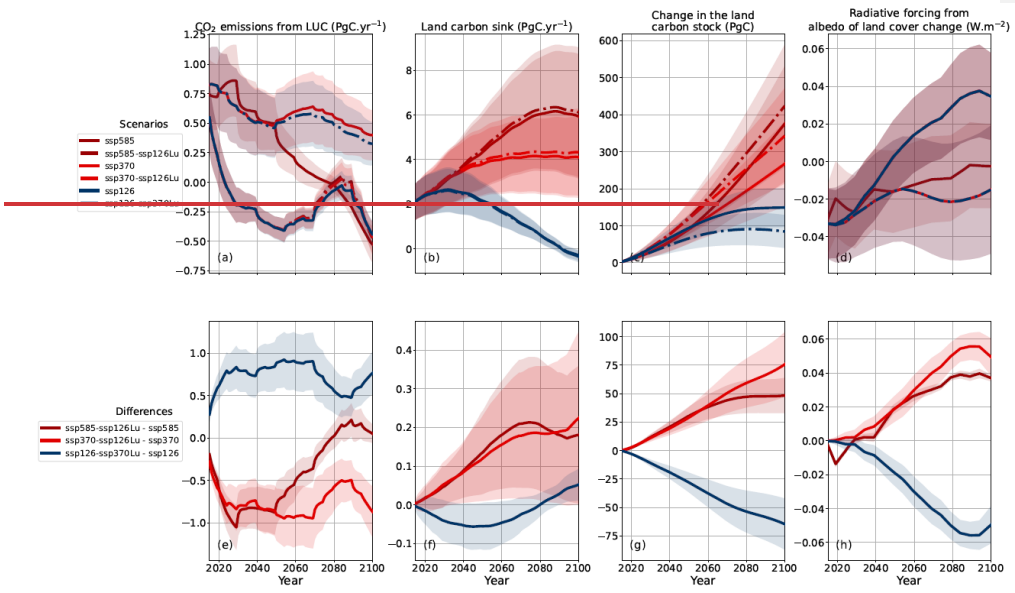

Figure A. 7: Effect of alternative land-use and land-cover change drivers in the scenarios *ssp126*, *ssp370* and *ssp585*. Here, the changes in the land carbon stock does not include the changes in the permafrost. The lines are the averages, and the shaded areas represent ± 1 standard deviation ranges.

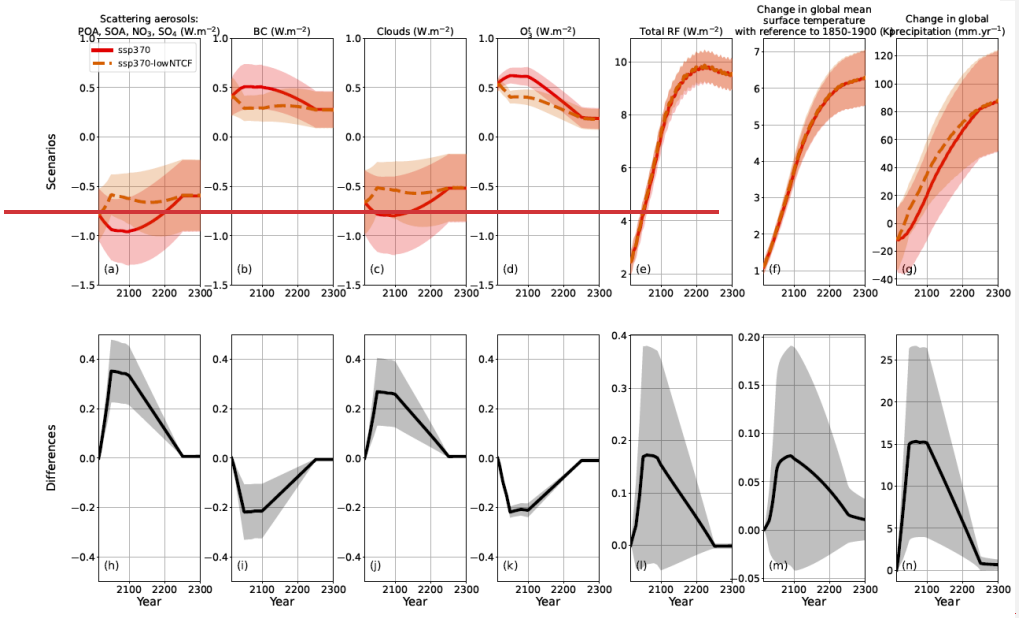

Figure A. 8: Effect of lower NTCF emissions in the SSP3-7.0. Extensions are shown only up to 2300. The lines are the averages, and the shaded areas represent ± 1 standard deviation ranges.

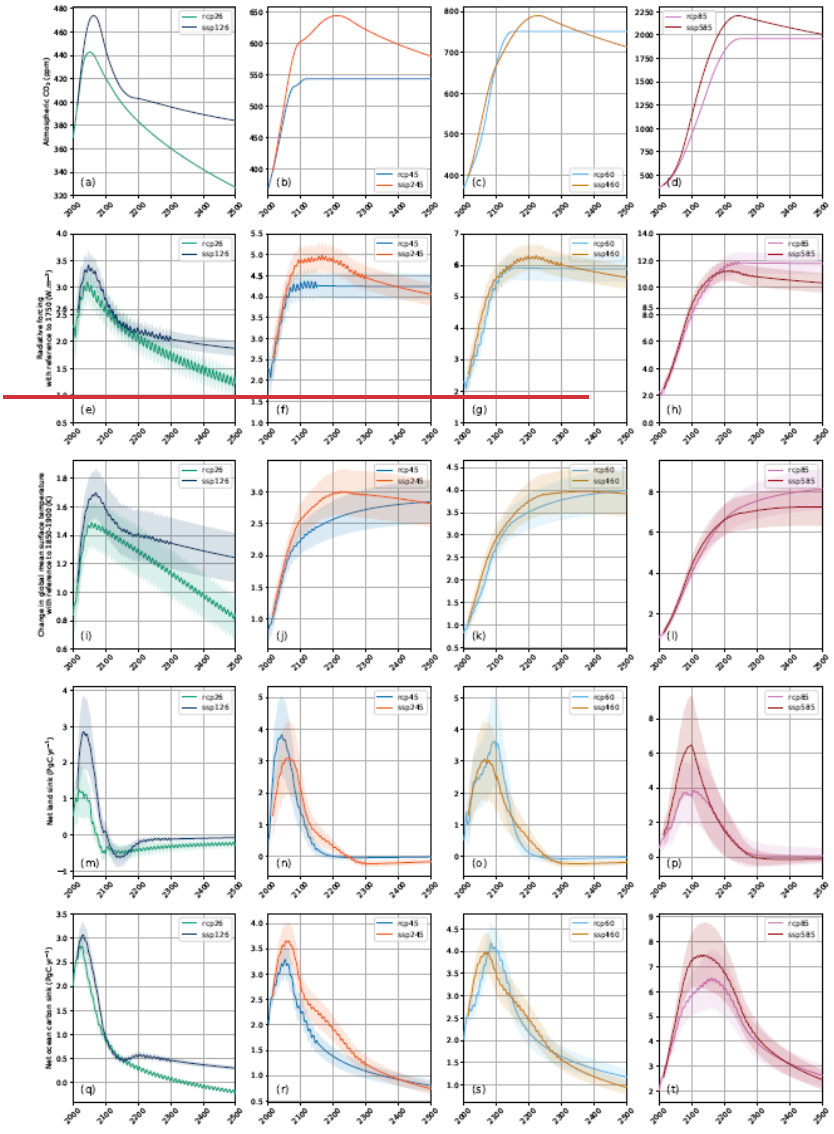
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
(k)
(l)
(m)
(n)
(o)
(p)
(q)
(r)
(s)
(t)
rcp26
ssp126
rcp45
ssp245
rcp60
ssp460
rcp85
ssp585

Figure A. 9: Comparison between RCPs (CMIP5) and SSPs (CMIP6). The lines are the averages, and the shaded areas represent ± 1 standard deviation ranges.

Appendix A: Method for excluding configurations and constraining outputs

In the exclusion round, we identify and discard the configurations that lead to a numerical divergence of the model as illustrated with Figure 2. Every experiment undergoes a thorough search, and we developed heuristic criteria to exclude these diverging runs by trial-error. We identify divergences occurring in high warming scenarios, mostly when the oceanic carbon sink drops and then oscillates. We explain this instability with the stratification of the ocean surface, as detailed in equation 4 of (Gasser et al., 2017). Some parametrizations under high warming scenarios exhibit an additional mode, not diverging in the strictest sense, yet, with the ocean carbon sink becoming a source and then switching back to a sink, which we identified as a physically unrealistic behaviour of the parametrization.

To discard the unrealistic configurations, we use the experiments *ssp585*, *ssp370*, *1pctCO2* and *abrupt-4xCO2* for their high warming over different timescales. We use the ocean sink, the land sink, the CO_2 emissions from LUC and the CO_2 emissions from permafrost, to ensure that the whole carbon cycle remains within reasonable boundaries. The criteria are set based on the performances of the remaining subset. In general, we use 20 PgC/yr in absolute value as a threshold for divergence. Over *ssp585* and *ssp370*, the domain is restrained to strictly positive values, due to the additional mode mentioned previously. Over *abrupt-4xCO2*, the criteria are applied over the last 50 years of the experiments only. In *1pctCO2*, the run is extended by another 100 years for better identification. As illustrated in Figure 2, mMost of the exclusions are related to ocean carbon sink, the other variables bring only little exclusions. TWe keep the 1118 configurations not causing any divergences in all the experiments are kept as a common set of configurations for all experiments.

The need for exclusion is stronger as the atmospheric concentration of CO_2 and the global surface temperature increase. We acknowledge that when a significant fraction of the configurations is excluded, confidence in our model's result is lowered, but such a limitation of the validity domain is inherent to reduced-complexity models. The model's results might as well depend on the set thresholds for exclusions. However, this bias is reduced through the constraining round because configurations with unrealistic carbon cycles receive a low likelihood.

We observed that in most cases, the reason of the exclusion is due to a diverging ocean sink. The ocean carbon cycle of OSCAR is its oldest module (Gasser et al., 2017), and should be redesigned for more stable behaviour under high-warming scenarios. A possibility is to increase the number of sub-timesteps in the oceanic carbon module to avoid this issue for a fraction of the configurations, but it comes at the expense of the computational cost of the model.

After this exclusion, the outputs of OSCAR are constrained using observations. As done for RCMIP phase 2 (Nicholls et al., 2021), the objective of this constraining round is to use the flexibility and the probabilistic frameworks of the reduced complexity models to synthesize lines of evidence with the modelling of the Earth system. With OSCAR, we assess the physical likelihood of the model's configurations using lines of evidence from the literature. For every constraint, we extend a method already used with OSCAR but with only one constraint (Gasser et al., 2020a; Le Quéré et al., 2018b). We assume a

Formatted: Font: Bold, Underline

Field Code Changed

Field Code Changed

distribution from which we derive the likelihood of every configuration, as illustrated in equation A1 of (Gasser et al., 2020a). The product of the probabilities over the set of constraints is the final likelihood of the configurations.

As the first observational constraint, we choose the surface air ocean blended temperature change over 2000-2019 with reference to 1961-1990 ~~are used~~, provided as an assessed range by RCMIP (Nicholls et al., 2021) from the HadCRUT 4.6.0.0 dataset (Morice et al., 2012). This constraint is meant to provide information on the climate system. To constrain the carbon cycle, we use compatible fossil fuel emissions. For now, OSCAR v3.1 is calibrated on CMIP5, which motivates the use of the compatible emissions of CMIP5, not those of CMIP6. An initial set of constraints based solely on observations had revealed that using projections helped the overall constraining round, thanks to the larger perturbation in the scenarios than in the historical period. Thus we choose the CMIP5 cumulative compatible fossil fuel emissions over the concentrations-driven historical and 4 RCPs ~~are used~~ (Ciais et al., 2013b). To further constrain the partitioning of the carbon sinks between land and ocean, we use data on the cumulative net ocean to atmosphere flux of CO_2 over 1750-2011 (Ciais et al., 2013b).

Formatted: Font: Bold, Underline

Formatted: Indent: First line: 0 cm

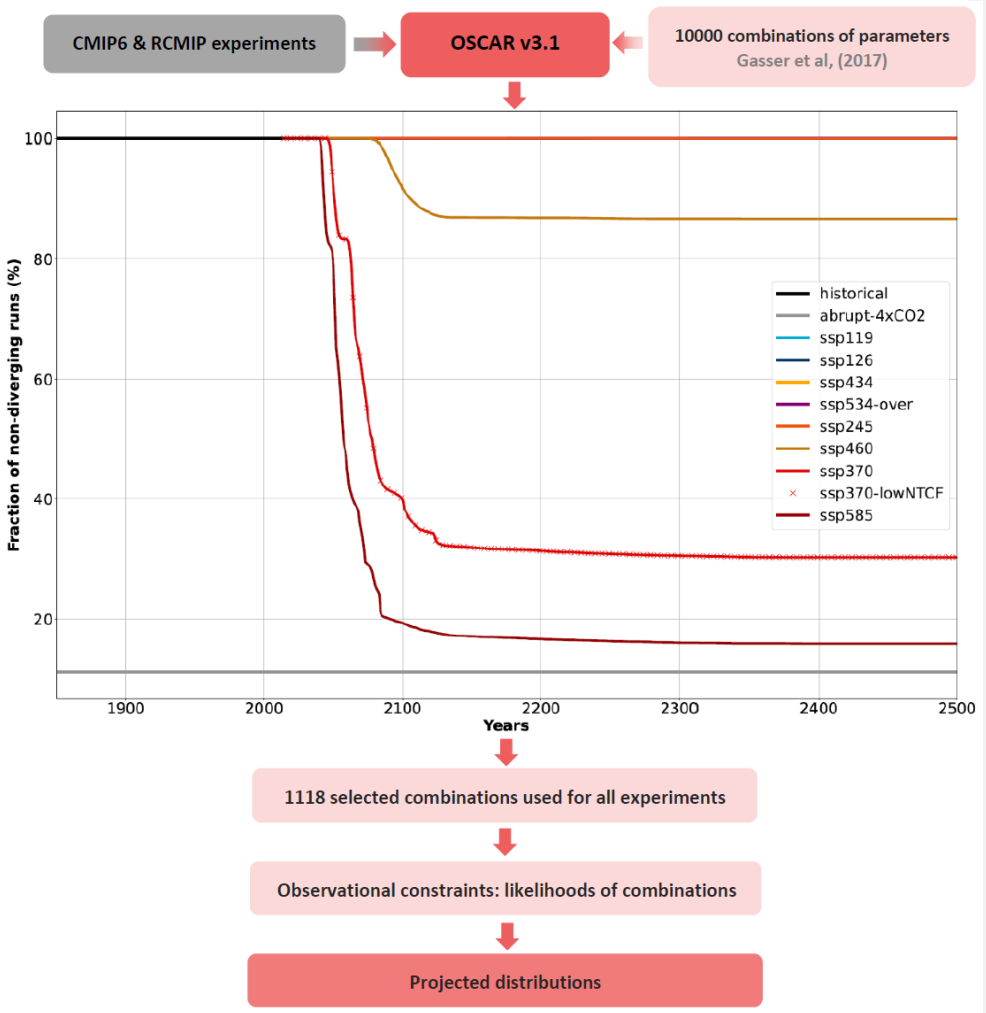

Figure A. 1: Conceptual description of the framework used in this study. The 10000 drawn configurations (Gasser et al., 2017) are used in OSCAR in a Monte-Carlo setup for all experiments. The exclusions are based on their exceedance to thresholds in the ocean sink,

land sink, CO_2 emissions form LUC and CO_2 emissions from permafrost. The remaining subset common to each experiment is then used for all. The likelihood of the kept configurations is then calculated **(Gasser et al., 2020a)** and applied to all experiments.

Appendix B: Behaviour of OSCAR

Formatted: Font: 10 pt, Bold, Underline

1. Carbon geoengineering

1.1. Idealized experiments

Experiments of CDRMIP are designed to investigate the consequences of carbon dioxide removal for the Earth system (Keller et al., 2018a). In *1pctCO2-cdr*, the atmospheric CO_2 increases by 1% every year (just like *1pctCO2*), but after 140 years, the atmospheric CO_2 decreases following a pathway at the same rath than the ramp-up period. Once CO_2 has returned to its preindustrial state, the experiment is extended over 1000 years. As shown in **Error! Reference source not found.**, the GSAT reaches 3.68 ± 0.39 K at the end of the ramp-up forcing, and it goes back to 0.85 ± 0.22 K at the end of the ramp-down forcing. For all variables, such as the CH_4 emissions from wetlands, removing CO_2 from the atmosphere during ramp-down effectively reduces the perturbation in the variable that was induced by the ramp-up, albeit within a different time frame that is typical of a dynamic hysteresis (Boucher et al., 2012). Once the global temperature change is sufficiently reduced, the permafrost carbon stock slowly reconstitutes itself as well. However, the whole Earth system is not fully recovered as soon as the preindustrial level of atmospheric CO_2 is reached. To return within 10% of the maximum perturbation at the end of the CO_2 ramp-up, it takes GSAT an average 110 extra years, and the land carbon stock an average 26 years. At the end of the 1000-year extension, the oceanic carbon stock remains at about 19% of its maximum perturbation.

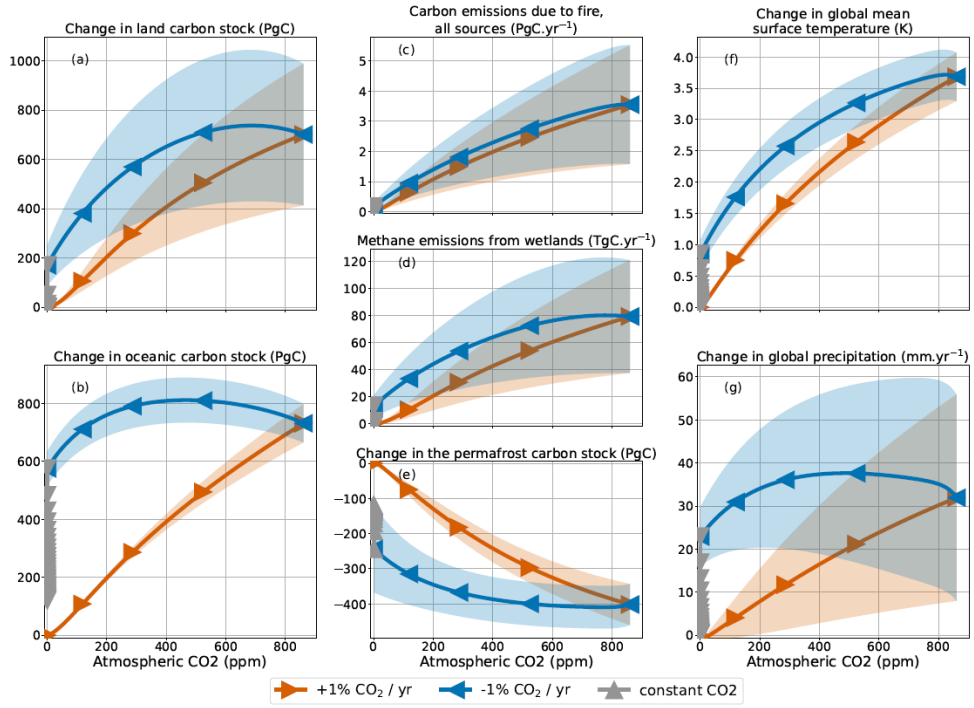

Figure B. 1: Reversibility experiment from CDRMIP. The orange lines correspond to the ramp-up of *1pctCO2-cdr*, the blue line to its ramp-down and the grey line to the 1000 years with constant atmospheric CO_2. The plain lines are the averages, and the shaded areas represent ± 1 standard deviation ranges.

Other CDRMIP experiments based on pulses of carbon emission or removal in an emission-driven configuration were performed to evaluate the response of the Earth system to CDR. These experiments are used to calculate the Absolute Global Warming and Temperature Potentials (AGWPs and AGTPs) of CO_2, which serves to establish the Global Warming and Temperature Potentials (GWPs and GTPs) of other greenhouse gases (Myhre et al., 2013). In *esm-pi-CO2pulse*, a 100 PgC pulse is emitted from the preindustrial environmental condition in 1860, whereas 100 PgC are removed in *esm-pi-cdr-pulse*. In *esm-yr2010CO2-CO2pulse*, the 100 PgC pulse is applied in 2015 but under 2010 environmental conditions, whereas these 100 PgC are removed at the same date in *esm-yr2010CO2-cdr-pulse*. We calculate timeseries of AGWPs and AGTPs under these experiments (). The differences to the reference pulse are shown in a different panel for clarity. We pinpoint that, just

like the other experiments, we are calculating these potentials with the interactive permafrost of OSCAR. The larger source of differences lies in the background: under preindustrial environmental conditions, emission pulses have a stronger AGWP or AGTP over the short term, but this is inverted over the longer term. Over the short term, this is due to the logarithmic expression of the CO_2 radiative forcing that is less saturated under preindustrial conditions. Over the long term, this is due to the deterioration of the carbon sink capacities under current conditions (Raupach et al., 2014). Similar reasons explain why a pulse of carbon removal cools the atmosphere slightly more over the short term than a pulse of emission warms it, but less over the long term. Our results cannot be compared to the final CDRMIP results yet, for they are unpublished, but they are consistent with those obtained with a model of intermediate complexity (Zickfeld et al., 2021).

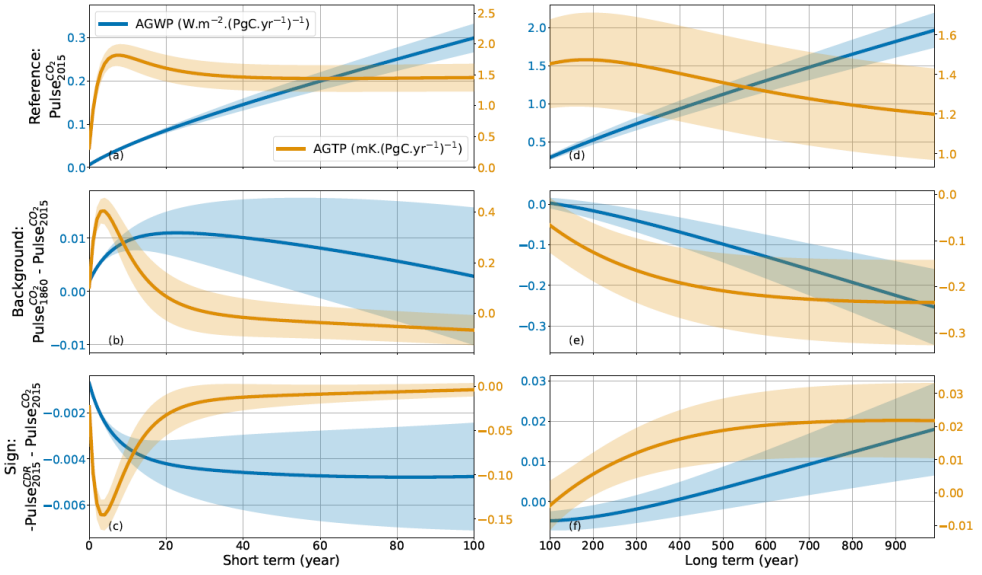

Figure B. 2: AGWP (blue) and AGTP (orange) of CO_2 for 100PgC of CO_2 emissions under actual environmental conditions. The dependency of this reference to a change of background is on the second line. The dependency to the sign of the pulse, emissions or removal, is on the third line. The lines are the averages, and the shaded areas represent ± 1 standard deviation ranges.

1.2. Alternative scenarios

The C4MIP (Jones et al., 2016) experiments *ssp534-over-bgc* and *ssp585-bgc* differ from *ssp534-over* and *ssp585* in that the prescribed CO_2 does not affect the total radiative forcing , thus causing a lower change in GSAT and maintaining a

relatively high carbon sinks efficiency. **Error! Reference source not found.** shows both carbon sinks under the variants and the base scenarios. Note that the *-bgc* experiments stem from a different historical simulation (*hist-bgc*). Under the high warming scenarios *ssp585*, climate change reduces the oceanic carbon sink by 1.93 ± 0.69 PgC.yr^{-1} and the net land carbon flux by 4.31 ± 1.93 PgC.yr^{-1} in 2100. Under the overshoot scenario *ssp534-over*, this difference is lower, owing to its declining atmospheric CO_2. Removing the impact of climate change on the carbon cycle increases the land carbon stock by 269 ± 52 PgC in *ssp534-over*, but by 501 ± 117 PgC in *ssp585* in 2100, due to the higher warming in the latter case. We note that the permafrost carbon stock drives most of the changes, because if permafrost is ignored in the *bgc* variant, these changes are reduced to 57 ± 32 PgC and 131 ± 77 PgC in *ssp534-over* and *ssp585* respectively.

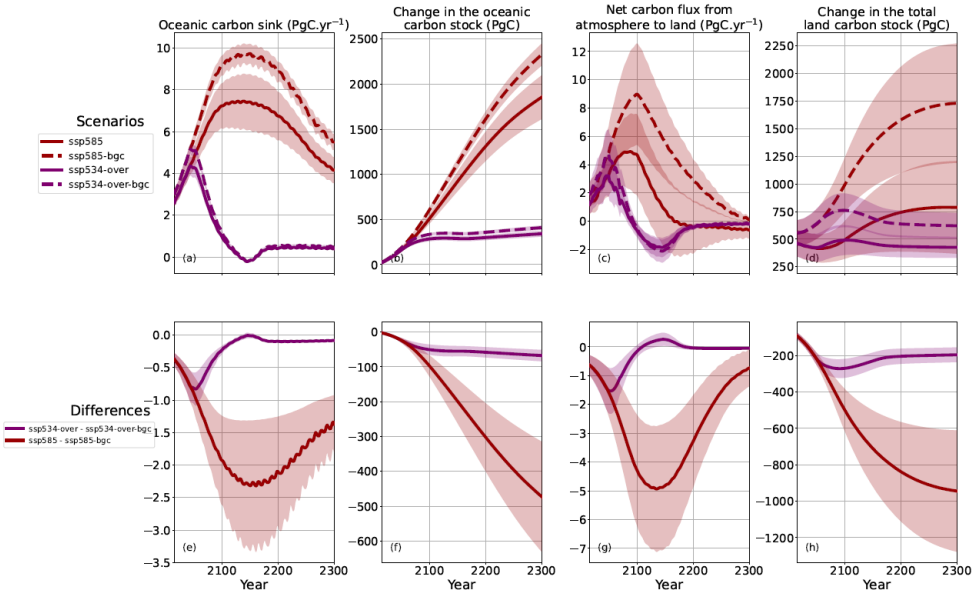

Figure B. 3: Effect of climate change on the carbon cycle in the scenarios *ssp534-over* and *ssp585*. The net flux from atmosphere from land is the sum of the land carbon sink, CO_2 emissions from land-use and land-cover change, and CO_2 and CH_4 emissions from permafrost. The changes in the total land carbon stock include those in the permafrost. Note that the increased uncertainty in the ocean sink before 2250 is an artefact of our exclusion procedure (see text on post-processing) that cannot capture the Monte Carlo members that already started diverging. Extensions are shown only up to 2300. The lines are the averages, and the shaded areas represent ± 1 standard deviation ranges.

2. Solar geoengineering

2.1. Idealized experiments

Experiments of GeoMIP (Kravitz et al., 2015) are designed to investigate the geoengineering techniques of Solar Radiation Management (SRM). Although OSCAR is not suited for all GeoMIP experiments, as it lacks any spatially resolved process, a few simulations remained accessible to our model. We run experiments *G1* and *G2*: G1 essentially follows *abrupt-4xCO2*, albeit with a changed incoming solar radiation that compensates for the radiative forcing caused by the increasing atmospheric CO_2. For *G2*, an identical principle is applied but using *1pctCO2* as a basis. As explained by (Kravitz et al., 2011), the change in solar radiation compensates solely for the radiative forcing of CO_2. However, it does not compensate for other radiative effects introduced by biogeochemical feedbacks, such as the fertilization by CO_2, affecting the carbon cycle, thus changing biomass burning emissions. **Error! Reference source not found.** shows that offsetting the CO_2 radiative forcing with a change in solar activity effectively compensates the change in GSAT. However, we simulate that the GSAT decreases in *G1* and *G2* to reach -0.08 ± 0.20 K and -0.07 ± 0.20 K, respectively, at the end of simulations. The compensation of the sole radiative forcing of CO_2 does not balance other feedbacks. There remains an additional radiative forcing, mostly due to changes in aerosols (as also shown in Figure), which results in this relatively small cooling in *G1* and *G2*. We estimate that in OSCAR about half of this effect is caused by the vegetation being fertilized by CO_2 and fuelling increased natural biomass burning emissions, and the remaining half is caused by the direct impact of GSAT on the atmospheric lifetime of aerosols (not shown). We note that the latter effect could be poorly estimated, in these specific experiments, as OSCAR's formulation for the lifetime of aerosols depends only on GSAT and not on the precipitation intensity.

Indeed, global precipitation does not respond in a similar way, because changes in atmospheric CO_2 and solar radiation have a different impact of the hydrological cycle (Andrews et al., 2010). In spite of a fully compensated GSAT change, global precipitation is significantly reduced in *G1* and *G2*, showing that such SRM technique does not entirely negate climate change. This demonstrates that OSCAR is capable of reproducing this well-established effect of this SRM technique (Boucher et al., 2013a). One added value of having a fully coupled ESM run these GeoMIP experiments is that we can also provide an estimate of the impact of the SRM technique on the carbon cycle. **Error! Reference source not found.** also shows that the land and ocean carbon stocks are increased in *G1* and *G2*, respectively by about 33% and 20% at the end of the simulations, owing to the loss of carbon sink efficiency that is avoided by maintaining the temperature to its preindustrial level.

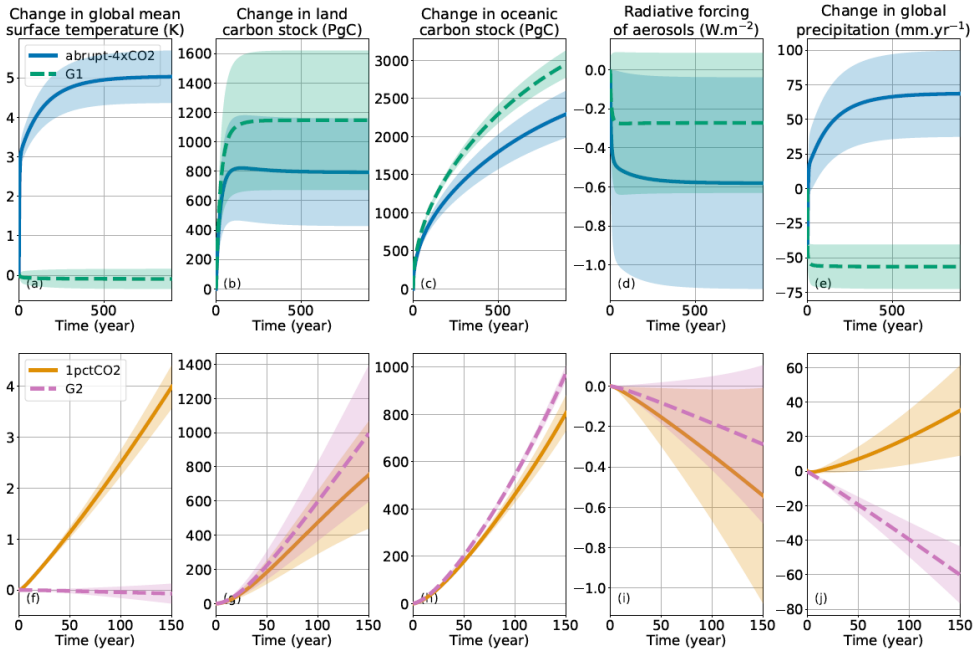

Figure B. 4: Experiments from GeoMIP compared to their DECK counterpart. The plain lines are the averages, and the shaded areas represent ± 1 standard deviation ranges.

2.2. Alternative scenarios

In addition to the few idealized experiments of GeoMIP (Kravitz et al., 2015) that are accessible to OSCAR, one scenario variant focusing on SRM was also feasible. The *G6solar* experiment stems from *ssp585*, but the solar constant changes from 2020 onwards to compensate the radiative forcing of *ssp585* and match the one of *ssp245*. As shown in **Error! Reference source not found.**, differences remain although the GSAT of *G6solar* decreases to a level comparable to *ssp245*. We calculate the change in solar constant as the difference from the radiative forcing of *ssp245* to *ssp585*. By construction, it excludes feedbacks caused by this change and does not fully cancel the change in global precipitation, just like in *G1* and *G2*. Consequently, the carbon stocks still increase in *G6solar*, even more than in *ssp585* thanks to the lower GSAT and despite lower global precipitation.

Formatted: Font: Italic

Formatted: Font: Italic

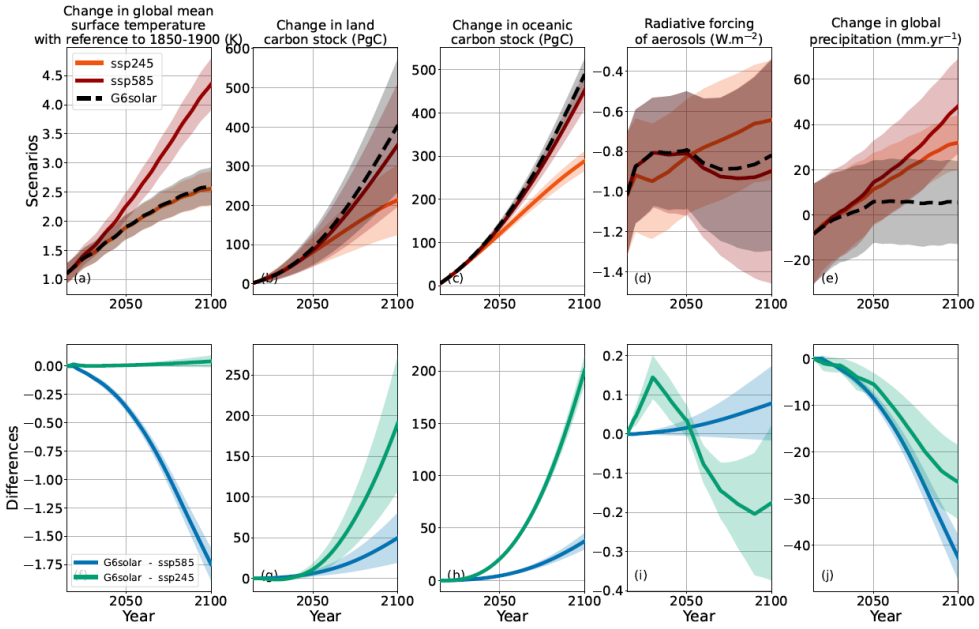

Figure B. 5: Effect of introducing SRM in the SSP5-8.5 to reach the SSP2-4.5. The lines are the averages, and the shaded areas represent ± 1 standard deviation ranges.

3. Land-use

3.1. Alternative historicals

LUMIP consists of experiments specifically focusing on land-use activities, and most of them are run by the Earth system models in a so-called "offline" fashion (Lawrence et al., 2016). It means that a reconstruction of past climate variables GSWP3 (Lawrence et al., 2016; van den Hurk et al., 2016) is prescribed to the model, so that the land module is actually decoupled from the rest of the model. Despite its simplicity, OSCAR has an added-value in running those simulations, as it embeds a book-keeping module that endogenously estimates CO_2 emissions from land-use and land-cover change. The main land carbon fluxes and stocks simulated under the reference experiment (dubbed *land-hist*) are shown in **Error! Reference source not found.**, along with three sets of sensitivity experiments described hereafter. The results are similar to those obtained recently with the same version of the model but with slightly differing forcings and a different constraint (Gasser et al., 2020b). The simulated land carbon stock decrease up to the 1970s, because of land-use activities emitting more CO_2 than the sink absorbs thanks to CO_2 fertilization and other factors. The carbon stock of 2010 is higher than the one of 1850 by only 1 ± 42 PgC. For comparison, the GCB 2020 provides for 1850-2014 a net budget for the land sink and CO_2 emissions from LUC of -5 ± 90 PgC (Friedlingstein et al., 2020).

The experiments *land-cCO2* and *land-cClim* are used to disentangle the contribution of CO_2 fertilization and changing climate on the land carbon cycle. In *land-cCO2*, the atmospheric CO_2 is constant and set to preindustrial value. In *land-cClim*, the climate drivers loop over the year 1901-1920 of the data set, thus simulating a preindustrial climate. **Error! Reference source not found.** shows the differences; for example, *land-hist – land-cCO2* illustrates the effect of atmospheric CO_2 on the variables of interest. Thanks to these experiments, we show that CO_2 is the main driver of the land sink in OSCAR, driving most of the trend, with climate bringing a significant interannual variability but virtually no trend, except over the recent past. In 2010, climate caused a small difference of -10 ± 10 PgC in total land carbon stock, while CO_2 did one of 141 ± 42 PgC. This has to be balanced with the results of the C4MIP idealized experiments where we saw OSCAR is less sensitive to climate change than CMIP5 models. Additionally, we see that the effect of climate and CO_2 on land-use and land-cover change emissions is minor, which is consistent with the fact that they are firstly determined by preindustrial carbon densities (Gasser et al., 2020b; Gasser and Ciais, 2013a).

A second set of experiments is meant to investigate the impact of land-use practices. Land-cover change contributed -152 ± 44 PgC to the 2010 change in land carbon stock since 1850, which corresponds to most of the total land-use and land-cover change emissions. Notably, it also reduced the land sink – an effect called the loss of additional sink capacity that has been diagnosed and quantified with OSCAR in the past (Gasser et al., 2020b; Le Quéré et al., 2018b; Gasser and Ciais, 2013a; Friedlingstein et al., 2020). Shifting cultivation (i.e. rapidly rotating land-use change between agriculture and natural ecosystems) had a relatively low impact on CO_2 emissions, leading to a change in land carbon stock of -8 ± 2 PgC at the end of the simulation in 2010. Similarly, wood harvest (in woody ecosystems that do not see land-cover change) had an overall impact of -16 ± 4 PgC. Both shifting cultivation and wood harvest have no impact at all on the land sink, by construction of

their formulation in OSCAR (Gasser et al., 2020b). Finally, the effect of having cropland-specific parameters in the model is isolated thanks to the *land-crop-grass* experiment, in which new croplands are treated as grasslands. Having grasslands instead of croplands increases both the land sink and the CO_2 emissions from land-use and land-cover change, resulting in a land carbon stock higher by 31 ± 26 PgC. All these values are entirely in line with an existing assessment of those land-use practices in which an earlier version of OSCAR took part (Arneth et al., 2017).

The third set of experiments relates to varying input data sets of land-use and land-cover change drivers. Two of these (*land-hist-altLu1* and *land-hist-alLu2*) relied on the two variations of the main LUH2 data set known as the "High" and "Low" variants (respectively) (Hurtt et al., 2020). We find that the so-called low variant leads to slightly higher land-use and land-cover change emissions amounting to a land carbon stock lower by 8 ± 2 PgC over the whole period. The high variant produces slightly lower total emissions, leading to a land carbon stock higher by 17 ± 5 PgC. Neither variant has a significant impact on the land sink. According to the description of these two variations (Hurtt et al., 2020), they differ from the default data set mostly in the harvest of biomass, and are very similar from 1920 onwards. The last LUMIP experiment run with OSCAR is one that uses the primary data set but an alternative starting year (*land-hist-altStartYear*). This required making an additional spin-up of the model under the environmental conditions and land cover of year 1700. Compared to the reference experiment, we find a slightly higher land sink after 1850 that decreases through time, owing to the ecosystems not being at steady state at that date. Similarly, emissions are slightly higher but the difference to the reference case tends towards zero as the legacy of land-use and land-cover change prior to 1850 fades away. The land carbon stock in 2010 is dominated by the increased land sink and amounts to a small increase of -17 ± 13 PgC in the land. Comparing the latter value with the total change in land carbon in the reference experiment suggests starting simulations in 1850 instead of 1700 or 1750 introduces a non-negligible bias in the CMIP6 exercise.

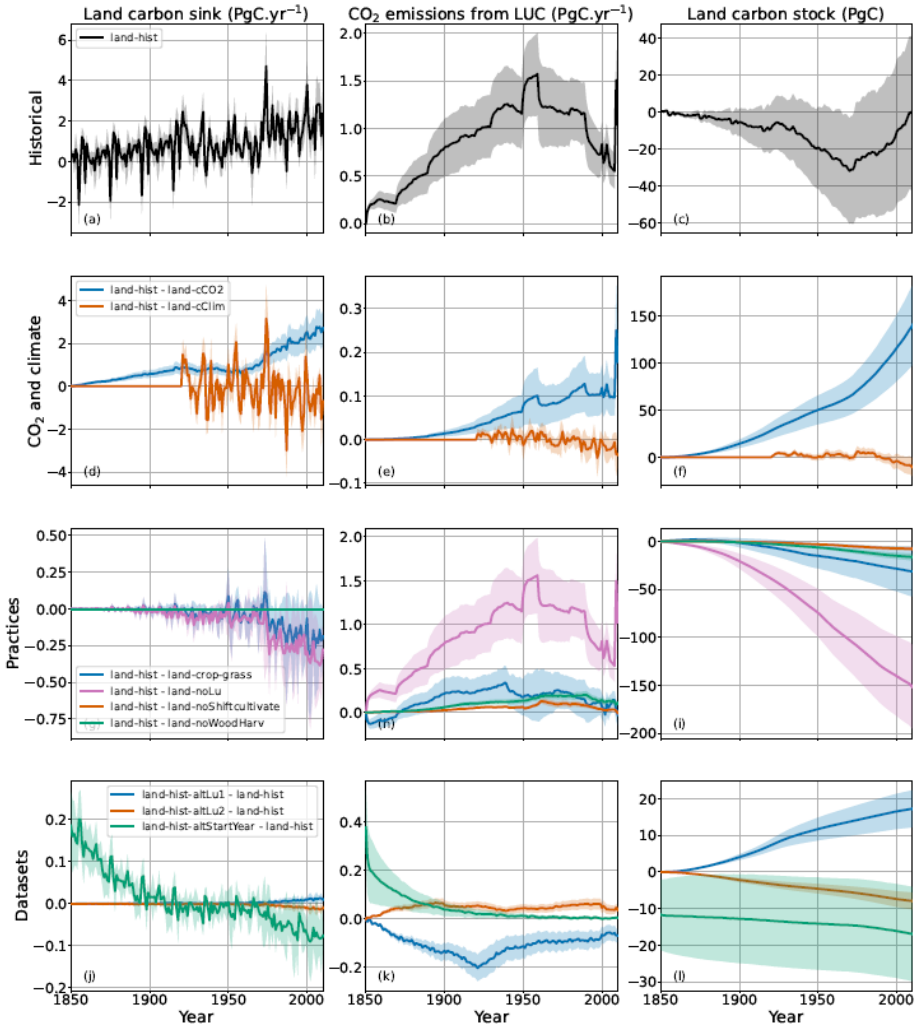

Figure B. 6: Land-use experiments from LUMIP. The first row of the figure corresponds to the reference experiment (*land-hist*) while other rows show sensitivity experiments as a difference to *land-hist*. *land-hist-altStartYear* is shown only from 1850 despite starting in 1700. The lines are the averages, and the shaded areas represent ± 1 standard deviation ranges.

Formatted: English (United States)

3.2. Alternative scenarios

LUMIP introduced variants of regular scenarios in which alternative land-use and land-cover change drivers coming from another scenario are prescribed (Lawrence et al., 2016), some of which being used in CDRMIP to assess afforestation (Keller et al., 2018b). Two such experiments are the pessimistic *ssp585* and *ssp370* combined with the land-use activities of the optimistic *ssp126* (named *ssp585-ssp126Lu* and *ssp370-ssp126Lu*, respectively). A third experiment consists in using the land-use of *ssp370* but under *ssp126*. (named *ssp126-ssp370Lu*). Comparisons of these experiments with their regular counterparts are shown in **Error! Reference source not found.**. As expected, changing the land-use scenario roughly replaces one SSP's land-use emissions by another's, albeit with some slight differences in the later stage of the simulations (i.e. after 2050) when atmospheric CO_2 and climate are significantly different from the reference scenario's, which has an impact in OSCAR because of transiently changing land carbon densities. The effect on the land carbon sink is also quantified, showing that sink capacity can be preserved by conserving natural ecosystems, although it remains a relatively small effect in absolute value. We note that the ability of properly isolating both effects (on land-use emissions and on the sink) is a specific feature of OSCAR that stems from the formulation of its land carbon cycle (Gasser et al., 2020b; Gasser and Ciais, 2013a), and we do not expect many complex ESMs to be able to provide such a partitioning. The overall effect on land carbon stock change in 2100 is 48 ± 15 PgC, 76 ± 28 PgC and -65 ± 23 PgC, in the *ssp585-ssp126Lu*, *ssp370-ssp126Lu* and *ssp126-ssp370Lu* scenarios respectively. While the land carbon stocks are affected, the change in land cover also affects the planetary albedo. The radiative forcing from albedo of land cover change are exchanged between *ssp126* and *ssp370*, but changes remain below 0.1 $W.m^{-2}$. The net combined effect on projected temperature cannot be estimated because these experiments are concentration-driven.

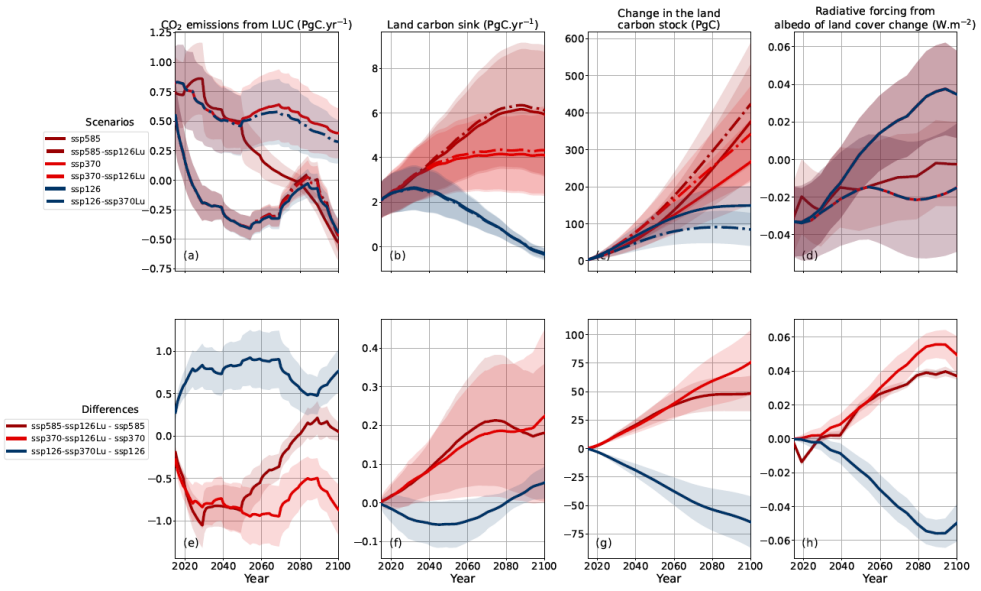

Figure B. 7: Effect of alternative land-use and land-cover change drivers in the scenarios *ssp126*, *ssp370* and *ssp585*. Here, the changes in the land carbon stock does not include the changes in the permafrost. The lines are the averages, and the shaded areas represent ± 1 standard deviation ranges.

4. Alternative scenarios for NTCFs

The *ssp370-lowNTCF* scenario is a variant of the *ssp370* differing by its lower emission of short-lived pollutants affecting the RF of NTCFs. As illustrated in **Error! Reference source not found.**, the variant leads to a somewhat equivalent warming, although with very slightly less cooling from NTCFs. This almost negligible effect on global temperature is actually the result of two large but compensating effects that manifest the most between 2050 and 2100. The lower emission of warming NTCFs leads to absorbing aerosols (i.e. BC) warming less by -0.21 ± 0.11 W m^{-2} and tropospheric ozone warming less by -0.21 ± 0.03 W m^{-2} in 2100. Conversely, it also leads to scattering aerosols cooling less by 0.33 ± 0.12 W m^{-2} and the indirect aerosol effects cooling less by 0.26 ± 0.13 W m^{-2} at the same date. This results in a small increase of the total radiative forcing of 0.15 ± 0.20 W m^{-2} and a GSAT change of only 0.07 ± 0.11 K. However, the difference in forcing agents between the two scenarios leads to a significant change in global precipitation that reaches 15 ±

11 mm yr^{-1} in 2100. The change in precipitation is consistent with our results for the GeoMIP experiments and what we know of the global water cycle (Shine et al., 2015).

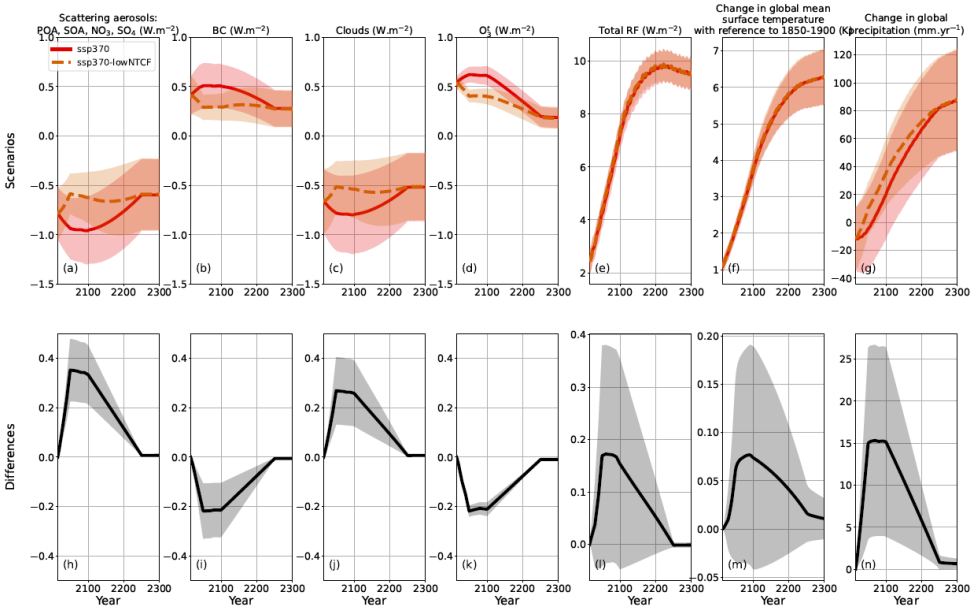

Figure B. 8: Effect of lower NTCF emissions in the SSP3-7.0. Extensions are shown only up to 2300. The lines are the averages, and the shaded areas represent ± 1 standard deviation ranges.

5. Comparison of two generations of scenarios

Initially, the SSPs scenarios were designed to reach the RF of RCPs in 2100, to provide a common grid for reading and comparing all the SSPs scenarios. Hence, the same four RF targets chosen in CMIP5 with the RCPs (2.6 W.m^{-2}, 4.5 W.m^{-2}, 6.0 W.m^{-2}, 8.5 W.m^{-2}) have also been chosen in CMIP6 with four out of the eight SSPs used. Yet, CMIP6 ESMs did not run RCPs, because these scenarios are not part of the CMIP6 experiments. Therefore, the difference between RCPs projections in CMIP5 and SSPs projections in CMIP6 under the same RF targets are due to both a change in the generation of ESMs and a change in scenarios. In **Error! Reference source not found.**, we represent both RCPs and SSPs under the same version of OSCAR, showing the difference due to the sole change in scenarios. These scenarios use different drivers, as illustrated with the atmospheric CO_2 prescribed to these concentration-driven experiments, usually with higher CO_2 concentrations in the

CMIP6 version. Except for the 8.5 target, the RF tends to be also higher in the CMIP6 version, compared to the CMIP5 version, meaning changes in other drivers are not enough to balance the CO_2 increase. While the 2.6 $W.m^{-2}$ and 8.5 $W.m^{-2}$ targets are reached in 2100, the 4.5 $W.m^{-2}$ and 6.0 $W.m^{-2}$ are not. However, our results can be compared to those of MAGICC in these two cases (IIASA, 2018a), and both reduced-complexity models are consistent. Because of the similar RF targets, GSAT are relatively similar over the 21st century, but RCPs and SSPs tend to dissociate later on. In 2300, moving from RCPs to SSPs changes GSAT by 18 ± 8%, 9 ± 3%, 5 ± 2% and -6 ± 1% in the four tested scenarios, respectively. Differences in other key variables such as the carbon sinks logically respond to these differences in atmospheric CO_2 and global temperature change, as also shown in **Error! Reference source not found.**.

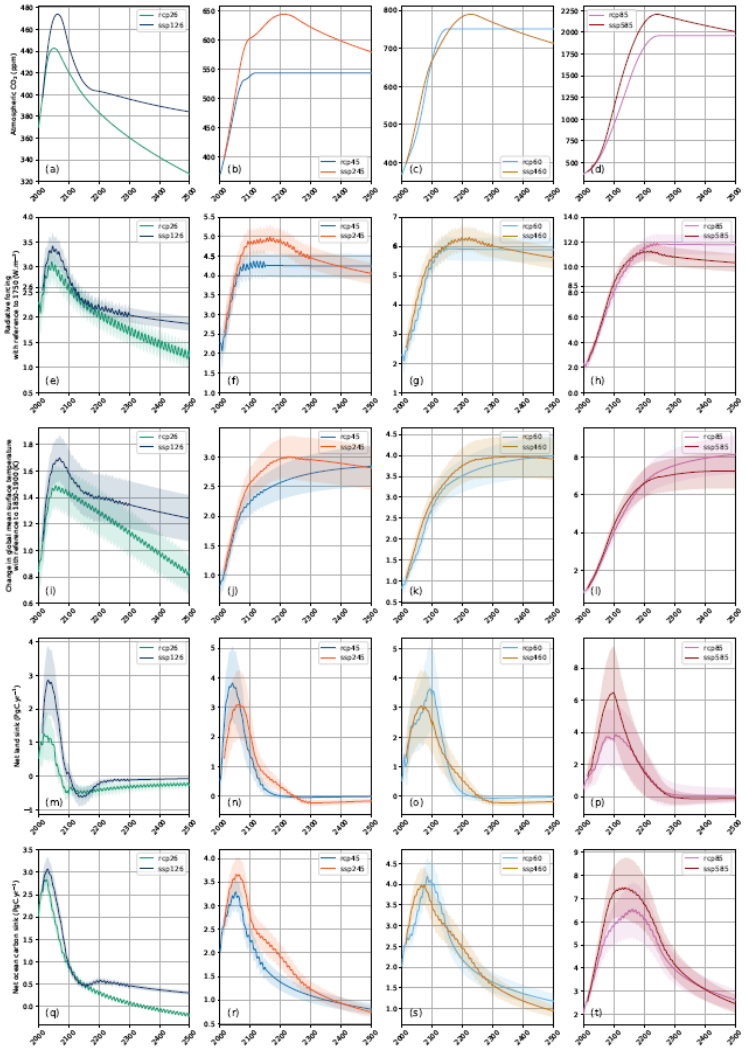

Figure B. 9: Comparison between RCPs (CMIP5) and SSPs (CMIP6). The lines are the averages, and the shaded areas represent ± 1 standard deviation ranges.

Formatted: Indent: First line: 0 cm