Universität Bremen

Constraining Uncertainties in Multi-Model Projections of the Future Climate with Observations

Doctoral Dissertation of

Manuel Schlund

March 2021

University of Bremen

Institute of Environmental Physics (IUP)

Constraining Uncertainties in Multi-Model Projections of the Future Climate with Observations

Doctoral Dissertation of

Manuel Schlund

A thesis submitted in fulfillment of the requirements for the degree
Doktor der Naturwissenschaften (Dr. rer. nat.)

First Supervisor: Prof. Dr. Veronika Eyring
Second Supervisor: Prof. Dr. Pierre Gentine

Submission Date: March 2021

Abstract

A precise quantification of climate change is crucial in order to assess optimal mitigation and adaptation strategies. Earth system models (ESMs), which are state-of-the-art climate models that allow numerical simulations of the complex physical, biological and chemical processes of the Earth system, are common tools to understand and project climate change. Due to the chaotic nature of the climate system, unknowns in future emission pathways and uncertainties in the climate models, projections of the future climate are associated with large uncertainties. The main focus of this thesis is the analysis of future climate projections from ESMs participating in the Coupled Model Intercomparison Project (CMIP) and the reduction of uncertainties in climate projections with observations.

In a first step, the climate sensitivity (i.e., the temperature response of the climate system to an external forcing) in the latest generation of ESMs from CMIP6 is evaluated. For the effective climate sensitivity (ECS), which is defined as the equilibrium temperature response that follows a doubling of the atmospheric carbon dioxide (CO_2) concentration, a multi-model mean (MMM) of 3.74 K and a multi-model range of 1.8–5.6 K are found. These values are higher than in any previous CMIP ensemble before. Moreover, a third of the analyzed CMIP6 models exceeds the upper bound of the likely ECS range of 1.5–4.5 K assessed by the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5) from 2013. Similarly, the transient response of the climate system to a doubling of CO_2, also known as transient climate response (TCR), shows an inter-model range of 1.3–3.0 K with an upper bound again higher than the likely range assessed in AR5 of 1.0–2.5 K. Possible reasons for the increased climate sensitivity in many CMIP6 models are aerosol-cloud interactions in the simulations due to changed aerosol schemes and changes in the microphysical representation of mixed-phase clouds over the Southern Ocean, which reduce the strong negative shortwave cloud phase change feedback in this region that is present in climate models from previous CMIP generations.

To reduce uncertainties in ECS projected by the CMIP6 models, eleven published emergent constraints on ECS (mostly derived from models participating in CMIP5, the predecessor generation of CMIP6) are systematically analyzed. Emergent constraints are potentially promising approaches to reduce uncertainties in climate model projections by combining observations and output from ESMs. The focus of this analysis is on testing if these emergent constraints hold for ESMs participating CMIP6. Since none of the emergent constraints considered here have been derived on the CMIP6 ensemble, the CMIP6 models can be used for cross-checking of the emergent constraints on a new model ensemble. The application of the emergent constraints to CMIP6 data shows a decrease in skill and statistical significance of the emergent

relationship for nearly all constraints, with this decrease being large in many cases. Consequently, the size of the constrained ECS ranges (66 % confidence intervals) widens by 51 % on average in CMIP6 compared to CMIP5. This is likely related to the increased multi-model spread of ECS in CMIP6, but may in some cases also be due to spurious statistical relationships or a too small number of models in the ensemble that the emergent constraint was originally derived from. The corresponding best estimates given by the emergent constraints also increase from CMIP5 to CMIP6 by 12 % on average. This can be at least partly explained by the increased number of high-ECS models in CMIP6 without a corresponding change in the constraint predictors, suggesting the emergence of new feedback processes rather than changes in strength of those previously dominant. The results support previous studies concluding that emergent constraints should be based on an independently verifiable physical mechanism and that process-based emergent constraints on ECS should rather be thought of as constraints for the process or feedback they are actually targeting.

To overcome these issues of single-process-oriented emergent constraints, an alternative approach based on machine learning (ML) is introduced. Since this new technique relies on a large number of data points in order to train the ML algorithm, the scalar climate sensitivity expressed as ECS or TCR is not an appropriate target variable. Therefore, gross primary production (GPP) as a process that contributes to climate sensitivity is studied as an alternative. GPP is the largest flux of the terrestrial carbon uptake and slows down global warming by removing CO_2 from the atmosphere. In this analysis, an existing emergent constraint on CO_2 fertilization is combined with a ML approach to constrain the spatial variations of multi-model GPP projections. In the first step of the two-step approach, observed changes in the CO_2 seasonal cycle at Cape Kumukahi, Hawaii are used to constrain the global mean GPP at the end of the 21st century (2091–2100) in Representative Concentration Pathway 8.5 simulations with ESMs participating in CMIP5 to (171 ± 12) GtC yr^{-1}, compared to the unconstrained model range 156–247 GtC yr^{-1}. In a second step, a ML model is used to constrain gridded future absolute GPP and gridded fractional GPP change in two independent approaches. For this, observational data is fed into the ML algorithm that has been trained on CMIP5 data to learn relationships between present-day physically-relevant diagnostics and the target variable. In a leave-one-model-out cross-validation approach, the ML model shows superior performance to the CMIP5 MMM. The new approach predicts a higher GPP increase in high latitudes and a lower GPP increase in regions closer to the equator.

Integrated Author's References

Parts of this thesis (text, figures and tables) are already published in the following peer-reviewed publications. More details on this are given in section 1.3 and in the beginning of the corresponding **??**.

Bock, L., Lauer, A., **Schlund**, **M.**, Barreiro, M., Bellouin, N., Jones, C., Meehl, G. A., Predoi, V., Roberts, M. J., & Eyring, V. (2020). Quantifying Progress Across Different CMIP Phases With the ESMValTool. *Journal of Geophysical Research: Atmospheres*, *125*(21). https://doi.org/10.1029/2019jd032321

Eyring, V., Bock, L., Lauer, A., Righi, M., **Schlund**, **M.**, Andela, B., Arnone, E., Bellprat, O., Brötz, B., Caron, L.-P., Carvalhais, N., Cionni, I., Cortesi, N., Crezee, B., Davin, E. L., Davini, P., Debeire, K., de Mora, L., Deser, C., . . . Zimmermann, K. (2020). Earth System Model Evaluation Tool (ESMValTool) v2.0 – an extended set of large-scale diagnostics for quasi-operational and comprehensive evaluation of Earth system models in CMIP. *Geoscientific Model Development*, *13*(7), 3383–3438. https://doi.org/10.5194/gmd-13-3383-2020

Lauer, A., Eyring, V., Bellprat, O., Bock, L., Gier, B. K., Hunter, A., Lorenz, R., Pérez-Zanón, N., Righi, M., **Schlund**, **M.**, Senftleben, D., Weigel, K., & Zechlau, S. (2020). Earth System Model Evaluation Tool (ESMValTool) v2.0 – diagnostics for emergent constraints and future projections from Earth system models in CMIP. *Geoscientific Model Development*, *13*(9), 4205–4228. https://doi.org/10.5194/gmd-13-4205-2020

Meehl, G. A., Senior, C. A., Eyring, V., Flato, G., Lamarque, J.-F., Stouffer, R. J., Taylor, K. E., & **Schlund**, **M.** (2020). Context for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models. *Science Advances*, *6*(26), eaba1981. https://doi.org/10.1126/sciadv.aba1981

Righi, M., Andela, B., Eyring, V., Lauer, A., Predoi, V., **Schlund**, **M.**, Vegas-Regidor, J., Bock, L., Brotz, B., de Mora, L., Diblen, F., Dreyer, L., Drost, N., Earnshaw, P., Hassler, B., Koldunov, N., Little, B., Tomas, S. L., & Zimmermann, K. (2020). Earth System Model Evaluation Tool (ESMValTool) v2.0-technical overview. *Geoscientific Model Development*, *13*(3), 1179–1199. https://doi.org/10.5194/gmd-13-1179-2020

Schlund, **M.**, Eyring, V., Camps-Valls, G., Friedlingstein, P., Gentine, P., & Reichstein, M. (2020a). Constraining Uncertainty in Projected Gross Primary Production With Machine Learning. *Journal of Geophysical Research: Biogeosciences*, *125*(11), e2019JG005619. https://doi.org/10.1029/2019jg005619

Schlund, **M.**, Lauer, A., Gentine, P., Sherwood, S. C., & Eyring, V. (2020b). Emergent constraints on equilibrium climate sensitivity in CMIP5: do they hold for CMIP6? *Earth System Dynamics*, *11*(4), 1233–1258. https://doi.org/10.5194/esd-11-1233-2020

Weigel, K., Bock, L., Gier, B. K., Lauer, A., Righi, M., **Schlund**, **M.**, Adeniyi, K., Andela, B., Arnone, E., Berg, P., Caron, L.-P., Cionni, I., Corti, S., Drost, N., Hunter, A., Lledó, L., Mohr, C. W., Paçal, A., Pérez-Zanón, N., ... Eyring, V. (2020). Earth System Model Evaluation Tool (ESMValTool) v2.0 – diagnostics for extreme events, regional and impact evaluation and analysis of Earth system models in CMIP. *Geoscientific Model Development Discussions*, *in review*, 1–43. https://doi.org/10.5194/gmd-2020-244

Contents

1 Introduction

1.1 Motivation

Climate change is one of the greatest challenges for humankind today. The warming of the climate system is "unequivocal" and "many of the observed changes are unprecedented over decades to millennia" (IPCC 2014). The changing climate increases the "likelihood of severe, pervasive and irreversible impacts for people, species and ecosystems" with "mostly negative impacts for biodiversity, ecosystem services and economic development" and amplifies "risks for livelihoods and for food and human security" (IPCC 2014). Potential drivers for climate change are all natural and anthropogenic substances and processes that may alter the Earth's energy budget. The human influence on the climate system is clear: Ever increasing emissions of greenhouse gases (GHGs) since the end of the pre-industrial era largely driven by economic and population growth led to atmospheric concentrations of carbon dioxide (CO_2), methane (CH_4) and nitrous oxide (N_2O) that are "unprecedented in at least the last 800000 years" (IPCC 2014). The effects of these GHG emissions and other anthropogenic drivers have been "detected throughout the climate system and are extremely likely to have been the dominant cause of the observed warming since the mid-20^{th} century" (IPCC 2014). The greenhouse effect is based on the optical properties of the GHGs: While the atmosphere is mostly transparent for the incoming solar (shortwave) radiation, the outgoing infrared (longwave) radiation that is reflected from the planet's surface is partly absorbed by the corresponding GHG molecules through an excitation of their corresponding vibrational modes and re-emitted isotropically. This process traps energy near the surface and leads to a warming of the Earth's surface and the lower atmosphere. The individual impact of the different drivers of climate change measured with the so-called *radiative forcing*, which quantifies the change in energy fluxes caused by changes in these drivers relative to pre-industrial conditions (IPCC 2013). The total radiative forcing is positive (which corresponds to a warming) and its largest contribution is caused by the increase in the atmospheric concentration of CO_2 due to fossil fuel emissions since the year 1750 (IPCC 2014). Apart from their physical warming effect on the climate, carbon-based GHGs like CO_2 also directly influence the global carbon cycle, an important biogeochemical cycle of the Earth. A crucial flux of the carbon cycle is the *gross primary production (GPP)*, which describes the carbon uptake of the terrestrial biosphere due to photosynthesis. Since this land carbon uptake absorbs about 30 % of the anthropogenic CO_2 emissions in today's climate (Friedlingstein et al. 2020), this process substantially slows down global warming and directly contributes to the magnitude of the climate sensitivity. Other important anthropogenic drivers of climate change are the emission of aerosols and land use/land cover changes. Apart from

anthropogenic drivers, there are also natural processes which impact the climate system like changes in the solar activity or volcanic eruptions. However, there is clear evidence that these natural drivers alone cannot explain the observed climate change (Haustein et al. 2017).

To successfully mitigate the massive impacts of climate change, a first important step is the understanding of climate change and its accurate quantification. Extremely valuable tools in this context are climate models, which allow us to simulate the behavior of the climate system under arbitrary conditions without having to perform (ethnically questionable) experiments in the real world. All around the world climate research institutes provide a variety of different climate models. Many of them participate in the Coupled Model Intercomparison Project (CMIP), which was initiated in 1995 by the Working Group on Coupled Modelling (WGCM) of the World Climate Research Programme (WCRP) to "better understand past, present and future climate changes arising from natural, unforced variability or in response to changes in radiative forcing in a multi-model context" (WCRP 2020). The CMIP models provide crucial input for the international climate assessments given by the Assessment Reports (ARs) of the Intergovernmental Panel on Climate Change (IPCC). For example, the latest generation of climate models from the most recent (sixth) phase of CMIP (Eyring et al. 2016; known as CMIP6) support the assessment of the upcoming IPCC Sixth Assessment Report (AR6), and their predecessor models from CMIP5 (Taylor et al. 2012) have been assessed as part of the Fifth Assessment Report (AR5) in 2013 (Flato et al. 2013). Modern-day climate models, which allow the simulation of biological and chemical processes in addition to the dynamics of the physical components of the Earth system, are also known as Earth system models (ESMs) and provide the most sophisticated simulations of the Earth's climate. In this thesis, the terms "climate model" and "Earth system model" are used interchangeably since most modern models participating in CMIP are ESMs or at least have ESM versions.

Simulations that extrapolate state of the climate system into the future are called *projections* of the future climate. These include idealized simulations with only a prescribed change in the atmospheric CO_2 concentration (e.g., an instantaneous doubling of CO_2 or a CO_2 increase of 1 % per year) as well as more realistic projections that consider different future scenarios (e.g., a fossil fuel–based future or a scenario that is based on a sustainable development). In many variables that are relevant for climate change, multi-model projections from CMIP show a large inter-model range (Collins et al. 2013; Flato et al. 2013). A crucial and policy-relevant example for this is the climate sensitivity, which refers to the change in the global mean near-surface air temperature (GSAT) that results from a change in the radiative forcing. Common metrics for this are the *effective climate sensitivity (ECS)*, which describes the equilibrium response of the climate system after a doubling of the atmospheric CO_2 concentration and the *transient climate response (TCR)*, which describes the transient response of the system to a CO_2 doubling. In AR5, both quantities have been assessed with large ranges of 1.5–4.5 K and 1.0–2.5 K for ECS and TCR, respectively (Stocker et al. 2013). The corresponding inter-model ranges from the CMIP5 models show similar results (Flato et al. 2013). For this reason, a careful statistical evaluation and further refinement of the output of multi-model climate projections is necessary in order to reduce associated uncertainties. A state-of-the-art technique for this is the

emergent constraints method, which uses a physically-based inter-model relationship between an observable quantity of the Earth system and a target variable to reduce uncertainties in the target variable with observations (Allen and Ingram 2002). An alternative approach is the weighting of climate models based on their performance (i.e., the distance of one model to observational products) and interdependence (i.e., the distance of one model to other climate models) (Knutti et al. 2017). These techniques form the baseline for the new analyses and results presented in this thesis, which partly utilize methods from a new emerging research field in climate sciences: artificial intelligence (AI) and machine learning (ML).

1.2 Key Science Questions

The aim of this thesis is to reduce uncertainties in multi-model climate projections with observations by addressing the following three key science questions:

1. What is the range of climate sensitivity in the latest generation of ESMs from CMIP6 compared to previous multi-model ensembles, and do we understand the processes that determine this uncertainty range?

2. Can uncertainties in climate sensitivity be reduced with observations using the emergent constraint approach?

3. Can uncertainties in multi-dimensional (gridded) climate projections be reduced with ML techniques and observations?

1.3 Structure of the Thesis

Parts of this thesis are already published in multiple peer-reviewed publications (two first-author studies and six co-author studies). A complete list of these is given on page vii. Wherever material from these studies is presented in this thesis, the pronoun "we" is used to increase readability by avoiding the passive voice and to acknowledge all involved contributors. However, unless stated otherwise, all contents from these publications (text, figures and tables) shown in this thesis originate from the author of this thesis. A detailed list of contributions to these studies is given in the corresponding **??**.

This thesis is structured as follows: **??** introduces the scientific background. This includes relevant literature that is used as a baseline for this thesis. **??** gives an overview over the contributions made to the Earth System Model Evaluation Tool (ESMValTool), an open-source software for the analysis of ESMs. These contributions helped improving the routine evaluation of ESMs which is useful for the entire scientific community and lead to co-authorship in four peer-reviewed studies (Eyring et al. 2020; Lauer et al. 2020; Righi et al. 2020; Weigel et al. 2020). Chapter 2 covers the assessment of climate sensitivity metrics like the ECS or TCR in the latest generation of ESMs from CMIP6. This work is already published in two

scientific publications (Bock et al. 2020; Meehl et al. 2020). Since the ECS and TCR are considerably higher in this new climate model generation, **??** describes the assessment of emergent constraints on the ECS for these CMIP6 models and compares these to results derived from CMIP5 models. The contents of this **??** are published in *Earth System Dynamics* (Schlund et al. 2020b). **??** focuses on a new method to reduce uncertainties in multi-dimensional (gridded) multi-model projections of the future climate with observations based on ML. As an example, the method is applied to GPP at the end of the 21st century, which is already published in the *Journal of Geophysical Research: Biogeosciences* (Schlund et al. 2020a). Finally, chapter 3 provides a summary of the results of this thesis and gives an outlook of possible future works.

2 Assessment of Climate Sensitivity in the CMIP6 Ensemble

In order to reduce uncertainties in multi-model climate projections, a first important step is the assessment of the desired target variables in the corresponding climate model ensemble. This is particularly relevant when a new generation of Earth system models (ESMs) is published that includes considerable modifications compared to the respective predecessor model versions. Currently, we are in a situation like this with the CMIP6 ensemble, for which new data is still released every day (as of March 2021). In light of the upcoming international climate assessment of the Intergovernmental Panel on Climate Change (IPCC)'s AR6, the evaluation of policy-relevant metrics like effective climate sensitivity (ECS) or transient climate response (TCR) for these new CMIP6 models is crucial since they provide vital information about the future climate of planet Earth. As the Earth System Model Evaluation Tool (ESMValTool) is an excellent tool that allows for a quick and robust evaluation of new Coupled Model Intercomparison Project (CMIP) data as soon as it gets published on the Earth System Grid Federation (ESGF) servers (see **??**), it is used here to assess and analyze ECS and TCR of the CMIP6 models and compare it to corresponding results of predecessor model generations and international climate assessments. This work is already published in two scientific papers: Bock et al. (2020) and Meehl et al. (2020). For Bock et al. (2020), the author of this thesis contributed their figure 8 (a bar chart showing the ECS for several CMIP3, CMIP5 and CMIP6 models), their figure 10 (map plots showing cloud-related feedback parameters for the CMIP5 and CMIP6 multi-model means (MMMs); see figure 2.4), code to create these figures with the ESMValTool and text for the manuscript (in particular a section about the evaluation of ECS). For Meehl et al. (2020), the author of this thesis contributed all figures and tables, code to produce these with the ESMValTool and text for the manuscript. This chapter first presents the evaluation of climate sensitivity for the CMIP5 and CMIP6 ensembles (section 2.1) and puts the corresponding ECS and TCR values into historical context (section 2.2). Finally, possible reasons for the apparent increase of climate sensitivity in the CMIP6 models are discussed in detail (section 2.3).

2.1 Evaluation of ECS and TCR in CMIP5 and CMIP6

Following **??**, ECS is calculated with the Gregory regression method using global mean near-surface air temperature (GSAT) and top of the atmosphere (TOA) net radiation data for 150 years of a 4xCO_2 simulation (Gregory et al. 2004). This calculation is illustrated in figure 2.1

for the CMIP6 MMM, which yields an ECS of 3.74 K when all 150 years of the run are used. However, similar to the CMIP5 ensemble, the exact value of ECS depends on the years considered in the Gregory regression (see **??**). Using only the first 20 years of the simulation gives a significantly lower ECS of 3.31 K than using only the last 130 years of the simulation, which gives an ECS of 4.05 K. As thoroughly describes in **??**, the reason for this is the state and time dependence of the climate feedback parameter, which is given by the slope of the Gregory regression line. Due to non-linear effects in the feedbacks, this slope changes over time, resulting in lower (higher) values of ECS when only early (late) years of the simulation are considered. In contrast to that, the transient response of the climate system TCR is calculated from GSAT data at the time of carbon dioxide (CO_2) doubling in a 1%CO_2 run (Bindoff et al. 2013; see **??**). Using these two definitions and the ESMValTool, ECS and TCR are assessed for all CMIP5 and CMIP6 models where the necessary temperature and radiation data is available (as of March 2021). Tables 2.1 and 2.2 give an overview over the results for the CMIP5 and CMIP6 models, respectively.

The first striking feature of these two tables is the increased MMM of ECS and TCR in CMIP6. For ECS, the CMIP6 MMM is about 16 % (0.51 K) higher than the corresponding CMIP5 MMM. For TCR, the relative difference between the two model ensembles is notably smaller with about 10 % (0.19 K). The spread in the multi-model ensembles (expressed as the multi-model standard deviation) shows an even larger increase in the CMIP6 ensemble: For ECS, the relative difference is about 49 % (0.72 K in CMIP5 to 1.07 K in CMIP6) and for TCR, the relative difference is about 17 % (0.36 K in CMIP5 to 0.42 K in CMIP6). The main reason for the increased MMM and spread in the CMIP6 ensemble is the existence of several models with very high values of ECS and TCR. In addition to that, there also exists a number of models with very low values of ECS. Therefore, the CMIP6 model ranges of ECS and TCR are well outside the corresponding assessed ranges given by the latest published Assessment Report (AR) of the IPCC from 2013 (Stocker et al. 2013) with 1.8–5.6 K for ECS (AR5: 1.5–4.5 K) and 1.3–3.0 K (AR5: 1.0–2.5 K) for TCR.

2.2 Comparison to Previous CMIP Generations and International Climate Assessments

To illustrate the results of the previous section and put them into historical context, figure 2.2 shows the assessed ECS and TCR ranges over the years from the Charney report (Charney et al. 1979) and the different ARs of the IPCC in combination with the corresponding modeled ranges from the different CMIP generations. Since the Charney report in 1979, the assessed range of ECS of 1.5–4.5 K has almost remained unchanged during the last 40 years (Albritton et al. 2001; Charney et al. 1979; Kattenberg et al. 1996; Mitchell et al. 1990; Stocker et al. 2013) with the exception of AR4, where the lower bound was temporarily increased to 2.0 K (Solomon et al. 2007). For AR2, AR4 and AR5, the corresponding climate model generations from CMIP1, CMIP3 and CMIP5 more or less agree with this assessed range (Flato et al. 2013;

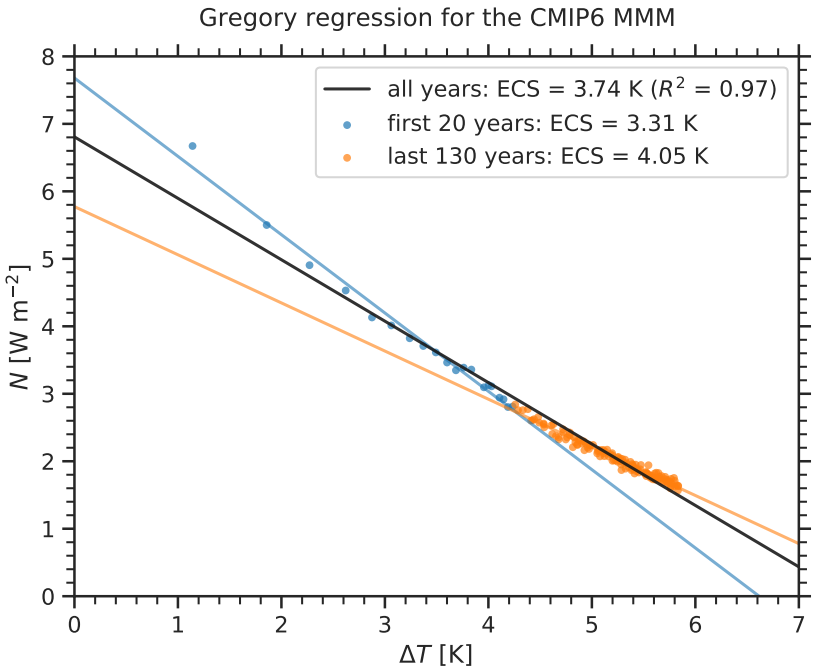

Figure 2.1: Gregory regression for the CMIP6 multi-model mean (MMM) following Gregory et al. (2004): Global and annual mean net top of the atmosphere radiation N versus the change in global and annual mean near-surface air temperature ΔT for 150 years of a simulation with an abrupt quadrupling of the atmospheric CO_2 concentration ($4xCO_2$) for the CMIP6 MMM (circles). To account for energy leakage and model drift, a linear fit of the corresponding pre-industrial control run is subtracted from the $4xCO_2$ simulation. A linear regression can be used to calculate the effective climate sensitivity (ECS) according to **??** with the radiative forcing F_{4x} (y-intercept) and the climate feedback parameter λ (slope) for all 150 years (black line), which results in ECS = 3.74 K. Due to non-linear effects, the exact value of ECS depends on the years considered for the Gregory regression (see **??** and **??**): Using only the first 20 years of the simulation (blue circles and line) yields a considerably smaller ECS than using only the last 130 years of the simulation (orange circles and line). Adapted with permission from Meehl et al. (2020).

Kattenberg et al. 1996; Randall et al. 2007). In contrast to that, some of the early climate models used in AR1 and some CMIP2 models used in AR3 exhibit values of ECS well above 4.5 K, resulting in a upper model range of about 0.6 K higher than the assessed range (Cubasch et al. 2001; Mitchell et al. 1990). However, these deviations from the assessed range are small compared to the ones present in the CMIP6 ensemble, which have become evident over the course of 2020. As shown in figure 2.2, both the lower and upper bound are more extreme than in any previous climate model generation, resulting in a CMIP6 ECS model range of 1.8–5.6 K (as of March 2021). This increase in range is particularly relevant for the upper bound, which is about 1.1 K higher than the assessed upper bound of 4.5 K. Moreover, it is not only a small fraction of models that exceeds this upper bound: out of the 42 analyzed CMIP6 models, a third (14 models) has an ECS above 4.5 K. In addition to the analysis presented here, several modeling centers independently confirmed high ECS values in their models (Andrews et al. 2019; Bodas-Salcedo et al. 2019; Gettelman et al. 2019; Wyser et al. 2020). A list of possible

Model	Index used in plots	ECS [K]	TCR [K]
ACCESS1-0	1	3.83	1.92
ACCESS1-3	2	3.53	1.63
BNU-ESM	3	3.92	2.49
CCSM4	4	2.94	1.67
CNRM-CM5	5	3.25	2.02
CNRM-CM5-2	6	3.44	1.79
CSIRO-Mk3-6-0	7	4.08	1.75
CanESM2	8	3.69	2.34
FGOALS-g2	9	3.38	1.38
GFDL-CM3	10	3.97	1.94
GFDL-ESM2G	11	2.39	1.12
GFDL-ESM2M	12	2.44	1.36
GISS-E2-H	13	2.31	1.74
GISS-E2-R	14	2.11	1.46
HadGEM2-ES	15	4.61	2.54
IPSL-CM5A-LR	16	4.13	2.00
IPSL-CM5A-MR	17	4.12	1.99
IPSL-CM5B-LR	18	2.60	1.53
MIROC-ESM	19	4.67	2.21
MIROC5	20	2.72	1.43
MPI-ESM-LR	21	3.63	2.03
MPI-ESM-MR	22	3.46	2.03
MPI-ESM-P	23	3.45	2.02
MRI-CGCM3	24	2.60	1.56
NorESM1-M	25	2.80	1.41
bcc-csm1-1	26	2.83	1.73
bcc-csm1-1-m	27	2.86	2.12
inmcm4	28	2.08	1.32
Multi-model mean		3.23	1.81
Multi-model standard deviation		0.72	0.36

Table 2.1: Effective climate sensitivity (ECS) and transient climate response (TCR) calculated for the CMIP5 models. Details on this are given in **??** and **??**, respectively. The multi-model mean (MMM) is calculated from the Gregory regression method using the MMM net top of the atmosphere (TOA) radiation and the MMM change in global mean near-surface air temperature (GSAT) similar to figure 2.1. The multi-model standard deviation is given by the sample standard deviation of ECS evaluated over all climate models (using the normalization $1/M$, where M is the number of models). Corresponding references for each model are given in **??**. Adapted with permission from Meehl et al. (2020).

reasons for the increased climate sensitivity in the CMIP6 ensemble including aerosol-cloud interactions and changes in the shortwave cloud feedback over the Southern Ocean is given in section 2.3.

In contrast to ECS, TCR has only been evaluated since AR1 in 1990, in which it is estimated with 2.3 K using only a single climate model (Bretherton et al. 1990; see single green circle in figure 2.2). Over the years, the modeled range of TCR has decreased from about 1.3–3.8 K in

Model	Index used in plots	ECS [K]	TCR [K]
ACCESS-CM2	29	4.72	2.10
ACCESS-ESM1-5	30	3.87	1.95
AWI-CM-1-1-MR	31	3.16	2.06
BCC-CSM2-MR	32	3.04	1.72
BCC-ESM1	33	3.26	1.77
CAMS-CSM1-0	34	2.29	1.73
CAS-ESM2-0	35	3.51	2.04
CESM2	36	5.16	2.06
CESM2-FV2	37	5.14	2.05
CESM2-WACCM	38	4.75	1.98
CESM2-WACCM-FV2	39	4.79	2.01
CMCC-CM2-SR5	40	3.52	2.09
CNRM-CM6-1	41	4.83	2.14
CNRM-CM6-1-HR	42	4.28	2.48
CNRM-ESM2-1	43	4.76	1.86
CanESM5	44	5.62	2.74
E3SM-1-0	45	5.32	2.99
EC-Earth3-Veg	46	4.31	2.62
FGOALS-f3-L	47	3.00	1.94
FGOALS-g3	48	2.88	
GISS-E2-1-G	49	2.72	1.79
GISS-E2-1-H	50	3.11	1.93
HadGEM3-GC31-LL	51	5.55	2.55
HadGEM3-GC31-MM	52	5.42	2.58
INM-CM4-8	53	1.83	1.33
INM-CM5-0	54	1.92	1.37
IPSL-CM6A-LR	55	4.56	2.32
KACE-1-0-G	56	4.48	1.41
MCM-UA-1-0	57	3.65	1.94
MIROC-ES2L	58	2.68	1.55
MIROC6	59	2.61	1.55
MPI-ESM-1-2-HAM	60	2.96	1.80
MPI-ESM1-2-HR	61	2.98	1.66
MPI-ESM1-2-LR	62	3.00	1.84
MRI-ESM2-0	63	3.15	1.64
NESM3	64	4.72	2.72
NorCPM1	65	3.05	1.56
NorESM2-LM	66	2.54	1.48
NorESM2-MM	67	2.50	1.33
SAM0-UNICON	68	3.72	2.27
TaiESM1	69	4.31	2.34
UKESM1-0-LL	70	5.34	2.79
Multi-model mean		3.74	2.00
Multi-model standard deviation		1.07	0.42

Table 2.2: As in table 2.1 but for the CMIP6 models. Corresponding references for each model are given in **??**. Adapted with permission from Meehl et al. (2020).

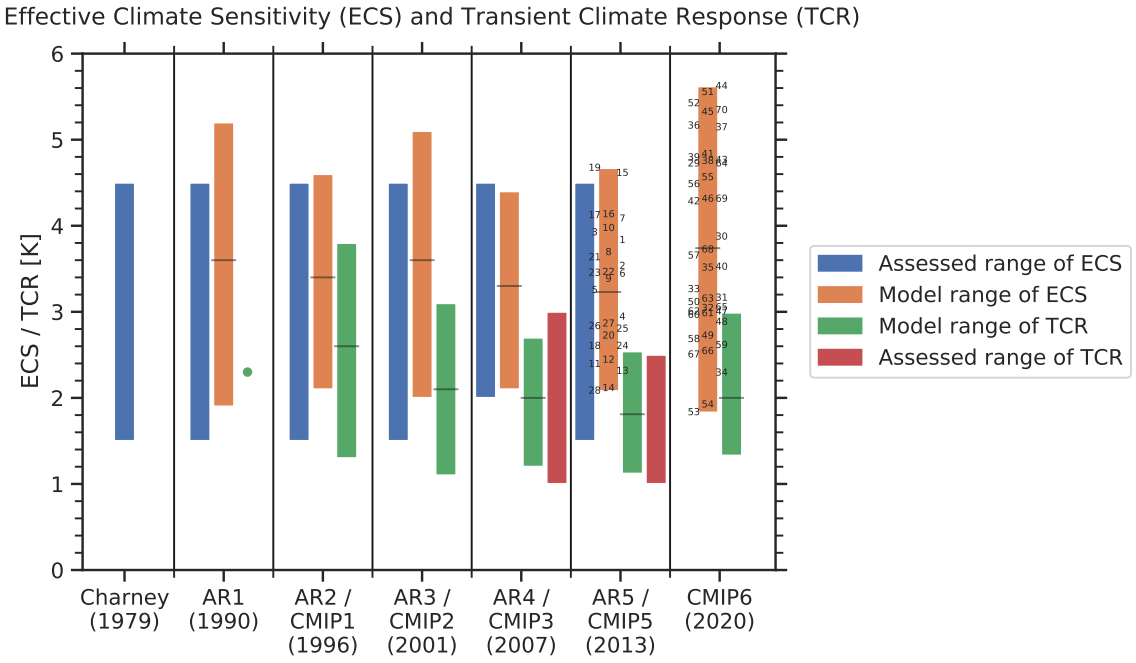

Figure 2.2: Assessed values of the effective climate sensitivity (ECS) (blue bars) and the transient climate response (TCR) (red bars) over the years in the Charney report from 1979 (Charney et al. 1979) and the subsequent Assessment Reports (ARs) of the Intergovernmental Panel on Climate Change (IPCC) (Albritton et al. 2001; Kattenberg et al. 1996; Mitchell et al. 1990; Solomon et al. 2007; Stocker et al. 2013). Orange and green bars represent the modeled ranges of ECS and TCR, respectively, from the different climate model ensembles of the Coupled Model Intercomparison Project (CMIP) with the corresponding multi-model mean (MMM) illustrated with the horizontal black lines (the green circle represents output from a single climate model). The numbers correspond to individual CMIP5 and CMIP6 models (see tables 2.1 and 2.2). Adapted with permission from Meehl et al. (2020).

the CMIP1 models (Kattenberg et al. 1996) to about 1.1–2.6 K in the CMIP5 ensemble (Flato et al. 2013). Similarly, the assessed range of TCR, which was first quoted in AR4 with 1.0–3.0 K (Solomon et al. 2007), has been reduced to 1.0–2.5 K in the subsequent AR5 (Stocker et al. 2013). While the corresponding climate model ranges from CMIP3 and CMIP5 agree with these assessed ranges (Flato et al. 2013; Randall et al. 2007), the upper bound of the CMIP6 model range (1.3–3.0 K) is well above the assessed upper bound of 2.5 K given by AR5. However, unlike ECS, also the lower bound of the CMIP6 model range increased compared to the CMIP5 ensemble, leading to a similar TCR model spread in both climate model generations. Due to the correlation of TCR and ECS given by **??**, the increase of TCR in many CMIP6 models is not surprising in light of the many high ECS models that are present the CMIP6 ensemble.

2.3 Possible Reasons for High Climate Sensitivity in CMIP6

Because of the massive political and societal relevance of ECS, its apparent increase in the CMIP6 ensemble is currently one of the most important questions for the entire climate

modeling community. For this reason, this section discusses possible reasons for the increased climate sensitivity in CMIP6.

Various improvements of the underlying physical, biological and chemical processes have been introduced to the CMIP6 ESMs in order to represent the coupled Earth system in more detail. Since many of these processes directly influence the models' ECS (Forster et al. 2020), determining reasons for the increased ECS in the CMIP6 ensemble is highly non-trivial. As discussed in detail by Meehl et al. (2020), one possible reason is the consideration of aerosol-cloud interactions. New prognostic aerosol schemes added to some CMIP6 models that include aerosol-cloud interactions could have produced overly large negative radiative forcing, which then required a stronger model response to greenhouse gas (GHG) forcing in order to reproduce the observed historical temperature trend. In fact, Meehl et al. (2020) find a weak inter-model relationship between the aerosol forcing and ECS for some CMIP6 models, with larger negative present-day aerosol forcing associated with higher ECS. Such relationships have also been found for previous climate model generations (Forster et al. 2013; Kiehl 2007). However, due to the varying aerosol forcing over time, these relations cannot be used to tune ECS based on the models' responses to aerosols over the different periods of the 20th century (Dittus et al. 2020). An example for a CMIP6 model with high ECS and a new aerosol scheme is the Hadley Centre Global Environmental Model (HadGEM) version 3. The new aerosol scheme in this model suppresses a strong negative cloud feedback over the Southern Ocean (Bodas-Salcedo et al. 2019), which is one of the most pristine aerosol regions in the world (Hamilton et al. 2014), where aerosols are predominantly of natural origin (e.g., marine emissions of gaseous dimethyl sulfide). Earlier HadGEM versions show a large reduction in cloud droplet size with warming over the Southern Ocean, which produces a negative shortwave cloud feedback. However, in HadGEM3, this negative feedback is completely suppressed or even positive due to a small increase of the cloud droplet size with warming, which eventually increases ECS due to an increased net cloud feedback (Bodas-Salcedo et al. 2019).

Another possible reason for the high climate sensitivity in the CMIP6 ensemble is related to cloud feedbacks (Bock et al. 2020). As already discussed in **??**, uncertainties in cloud feedbacks are a major source of uncertainty of ECS in modern ESMs (Boucher et al. 2013; Flato et al. 2013). Thus, changes in processes that are related to cloud feedbacks immediately impact the climate models' climate sensitivity. The particular feedback mechanism that is thought to be relevant for the high ECS in CMIP6 is connected to cloud phase changes over the Southern Ocean that are present in earlier CMIP generations (e.g., CMIP5). This so-called *cloud phase change feedback* is illustrated in figure 2.3 (green arrows). As the climate warms, the predominantly ice clouds over the Southern Ocean in these climate models become liquid clouds (McCoy et al. 2015). As a result, the cloud will get more reflective to the incoming solar radiation since a cloud consisting of smaller liquid droplets reflects more sunlight than a cloud consisting of larger ice crystals (for a fixed water content), which has a cooling effect on the climate system. In addition, predominantly liquid clouds tend to precipitate less than mixed-phase clouds (composed of liquid water and ice). This leads to a higher liquid water content of the

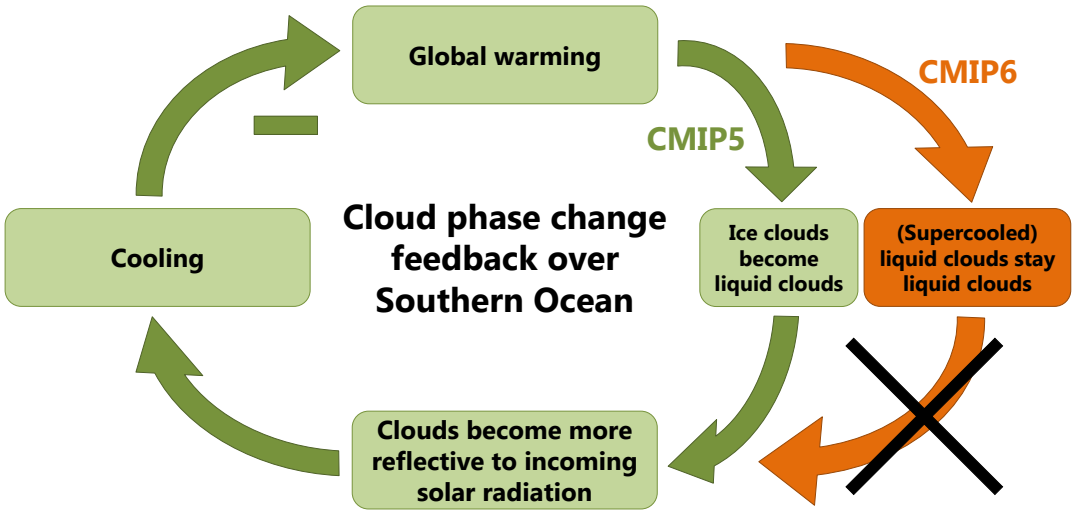

Figure 2.3: Schematic illustration of the strong negative shortwave cloud phase change feedback over the Southern Ocean, which is present in earlier CMIP model generations like CMIP5 (green arrows). Due to substantial improvements in the microphysical representation of mixed-phase clouds in many CMIP6 models, the predominantly ice clouds over the Southern Ocean in the present-day climate of previous climate model generations have been replaced with clouds that predominantly consist of (supercooled) liquid clouds (Bodas-Salcedo et al. 2019; Gettelman et al. 2019). Thus, in the affected CMIP6 models, this cloud phase change due to warming is reduced by allowing for supercooled cloud liquid, which substantially reduces the strong negative cloud phase change feedback (orange arrows). This leads to an increase in the global net cloud feedback (Bodas-Salcedo et al. 2019; Tan et al. 2016) and the effective climate sensitivity (ECS) in the corresponding CMIP6 models (Andrews et al. 2019; Gettelman et al. 2019), which is a possible explanation for the high ECS values found in many CMIP6 models.

cloud and a further amplification of the cooling effect. Overall, this forms a strong negative shortwave cloud feedback over the Southern Ocean in climate models from previous CMIP generations.

However, it is well known that the low-level mixed-phase clouds over the Southern Ocean in these earlier climate models are biased towards too large amounts of ice crystals and too little amounts of supercooled liquid water when compared to satellite observations (Bodas-Salcedo et al. 2016). Thus, the size of the resulting strong negative cloud phase change feedback in these models has long been questioned (McCoy et al. 2015; Tan et al. 2016). Improvements in the representation of mixed-phase clouds are known to reduce this long-standing cloud phase bias over the Southern Ocean (Bodas-Salcedo et al. 2016; McCoy et al. 2016) and to improve the representation of both cloud microphysical structure and cloud radiative impacts in this region (Hyder et al. 2018; Kay et al. 2016). For this reason, the microphysical representation of mixed-phase clouds has been updated in many CMIP6 models. For example, in CESM2, a new mixed-phase ice nucleation scheme by Hoose et al. (2010) replaces the Meyers et al. (1992) empirical scheme used in predecessor models (Gettelman et al. 2019). Simulations with the Meyers et al. (1992) scheme show almost no supercooled liquid water in high-latitude clouds and a strong

negative cloud phase feedback when ice turns to liquid. On the contrary, simulations with the new ice nucleation scheme by Hoose et al. (2010) show a strong reduction in this feedback and are more consistent with observations of the current climate (Gettelman et al. 2019). Another example is the HadGEM3 model, for which a new cloud scheme that accounts for turbulent production of liquid water in mixed-phase clouds has been implemented (Bodas-Salcedo et al. 2019). The new scheme parameterizes the role of subgrid-scale turbulence in the production and maintenance of supercooled liquid water (Furtado et al. 2016). As a consequence, the cloud liquid water content over the Southern ocean is increased, which is in closer agreement to observations compared to earlier HadGEM versions. Moreover, the increase in the cloud liquid water content with warming is reduced in HadGEM3, leading to a reduction in the negative cloud phase change feedback over the Southern Ocean (Bodas-Salcedo et al. 2019). By evaluating short-term forecasts created with HadGEM3, Williams et al. (2020) additionally show that the new climate model version improves in traditional metrics used for numerical weather prediction, which indicates an improvement in the physical processes simulated by the climate model.

Overall, these changes in many of the CMIP6 models lead to a better agreement of the CMIP6 MMM of the TOA shortwave cloud radiative effect (CRE) with corresponding observations (Bock et al. 2020). Moreover, the strong negative cloud feedback described above that results from a cloud phase change from ice clouds in the present-day to liquid clouds in the future is substantially reduced in those CMIP6 models that simulate predominantly liquid clouds over the present-day Southern Ocean. This is illustrated in figure 2.3 (orange arrows). As a consequence, this reduction of the negative shortwave cloud phase change feedback increases the global net cloud feedback (Bodas-Salcedo et al. 2019; Tan et al. 2016) and eventually the ECS of the affected CMIP6 models (Andrews et al. 2019; Gettelman et al. 2019), offering a possible explanation for the high ECS values found in many CMIP6 models.

To further confirm this hypothesis, the global distributions of relevant cloud feedback parameters (expressed by the net, shortwave and longwave CRE feedback parameters; see **??**) are shown for the CMIP6 MMM and its difference to the CMIP5 MMM in figure 2.4. First of all, the comparison of figures 2.4a, 2.4c and 2.4e demonstrates that the shortwave component dominates the net CRE feedback parameter over large swaths of the globe in the CMIP6 MMM. The sign change at approximately 60 °S in the shortwave CRE feedback parameter (see figure 2.4c) roughly shows where the climate models simulate a high fraction of ice clouds (south of 60 °S) for which possible phase changes under global warming are potentially important. Such possible phase changes (from ice to liquid) contribute to a negative shortwave cloud feedback in this region as liquid clouds tend to be more reflective than ice clouds with a similar cloud water content. Figures 2.4b and 2.4d show a higher shortwave and net CRE feedback parameter over the Southern Ocean in CMIP6 compared to CMIP5. This is consistent to the hypothesis presented in the paragraph above, which claims that the reduced negative cloud phase change feedback over the Southern Ocean in CMIP6 leads to an increased net cloud feedback parameter in comparison to older CMIP generations. This result is further supported by Zelinka et al. (2020), who find a larger positive zonal mean net cloud feedback

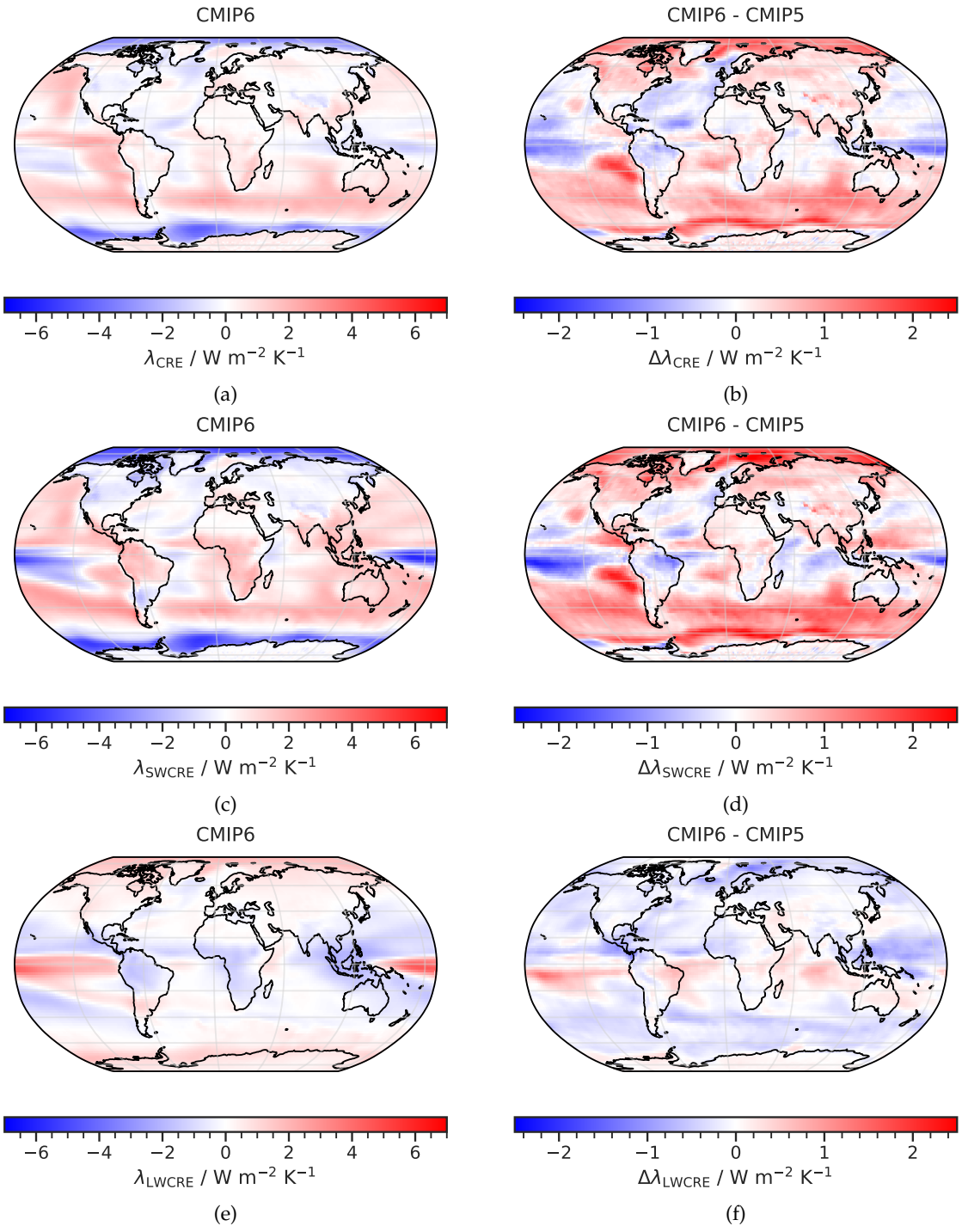

Figure 2.4: Geographical distributions of the net (a), shortwave (c) and longwave (e) cloud radiative effect (CRE) feedback parameters for the CMIP6 multi-model mean (MMM) and the corresponding differences to the CMIP5 MMM, again for the net (b), shortwave (d) and longwave (f) components. The CRE feedback parameters for each grid cell are calculated with a Gregory regression of the grid cell's CRE versus the change in the global mean near-surface air temperature (see **??**). Adapted with permission from Bock et al. (2020).

parameter in CMIP6 compared to CMIP5 due to a more positive (reduced negative) low-level cloud feedback, primarily in the extratropics. In comparison to earlier CMIP generations, the CMIP6 models exhibit weaker increases in the extratropical low-level cloud cover and the corresponding liquid water content as a result of global warming. Moreover, Zelinka et al. (2020) show that this mainly results from an increase in the liquid condensate fraction in the CMIP6 clouds for the pre-industrial and present-day periods, leading to the aforementioned reduction in the negative cloud phase change feedback on warming and eventually to higher ECS values.

3 Conclusion

3.1 Overall Summary

The analysis of future climate projections from numerical climate model simulations is of paramount importance to assess future climate change under different forcing scenarios. Since this involves metrics of public interest like the allowable fossil fuel emissions to meet particular warming targets, for example the 1.5 °C of the Paris Agreement (UNFCCC 2015), research in this field of science is not only relevant to climate scientists but to policymakers and the whole human society. In the light of the large spread of climate sensitivity in the most recent generation of climate models from CMIP6 (Meehl et al. 2020), a careful statistical evaluation and refinement of the output of multi-model climate projections is as relevant as ever. This thesis quantifies the associated uncertainties, presents the evaluation of established methods to reduce these uncertainties in the new climate model ensemble and finally introduces a novel alternative technique based on supervised machine learning (ML).

To answer the key science questions posed in section 1.2 and to ensure a consistent evaluation of the participating climate model ensembles in this thesis, the ESMValTool is used, an open-source community diagnostics and performance metrics tool for the routine evaluation of ESMs. All analyses shown in this thesis are performed with the ESMValTool. Apart from that, further substantial changes and additions to the code base of the tool have been implemented as part of this thesis (see **??**), which led to co-authorship in the technical and scientific documentation of the ESMValTool (Eyring et al. 2020; Lauer et al. 2020; Righi et al. 2020; Weigel et al. 2020). Since the open-source tool is freely available, the code that has been implemented as part of this thesis is beneficial for the entire scientific community.

In the first study of this thesis, the climate sensitivity metrics ECS and TCR are evaluated for the latest generation of ESMs from CMIP6. This work, which is presented in chapter 2 and already published in Bock et al. (2020) and Meehl et al. (2020), directly addresses key science question 1 ("*What is the range of climate sensitivity in the latest generation of ESMs from CMIP6 compared to previous multi-model ensembles, and do we understand the processes that determine this uncertainty range?*"). For ECS, a CMIP6 model range of 1.8–5.6 K is found, which is higher on the upper and lower end than any model range from previous CMIP generations before. In comparison to CMIP5, the CMIP6 MMM of ECS is about 16 % higher (3.74 K in CMIP6 versus 3.23 K in CMIP5). Moreover, especially the upper bound of the CMIP6 models range is considerably larger than in the assessed range of 1.5–4.5 K given by the latest published AR of the IPCC from 2013 (Stocker et al. 2013). The assessed upper bound of 4.5 K is exceeded by a third of the CMIP6 models, with many models showing ECS values above 5 K. For

TCR, the model range of CMIP6 is 1.3–3.0 K, which also exceeds the CMIP5 range of 1.1–2.5 K and the assessed range from AR5 of 1.0–2.5 K. One possible reason for the increased climate sensitivity in many CMIP6 models is the addition of prognostic aerosol schemes that include aerosol-cloud interactions which could have produced overly large negative radiative forcing in some models. To correctly reproduce the observed historical temperature trend, a stronger model response to GHG forcing is required, resulting in a higher climate sensitivity. A further reason for the high climate sensitivity in CMIP6 is a change in the microphysical representation of mixed-phase clouds over the Southern Ocean in some models. This change was mainly implemented to address a long-standing cloud phase bias over the Southern Ocean in the climate models, which has been successfully reduced. Moreover, the CMIP6 MMM shows an improved simulation of the shortwave CRE when compared to observations. However, the change in the representation of mixed-phase clouds also substantially reduces the strong negative shortwave cloud feedback over the Southern Ocean that is present in previous CMIP generations and that results from a cloud phase change from ice clouds in the present-day to liquid clouds in the future. In the affected CMIP6 models, this cloud phase change due to warming is reduced since these models simulate less cloud ice over the present-day Southern Ocean than their predecessor versions with no supercooled cloud liquid.

In order to reduce this large range of ECS in the latest generation of climate models, already-published emergent constraints that have been derived on models from the previous CMIP generations CMIP3 and CMIP5 are evaluated for their skill in the CMIP6 ensemble. Emergent constraints use a physically-based inter-model relationship between an observable quantity of the Earth system and a target variable to reduce uncertainties in the target variable with observations (Allen and Ingram 2002). In total eleven emergent constraints on ECS are assessed, which are mostly related to cloud feedbacks since these constitute the most important source of uncertainty for ECS (Boucher et al. 2013; Flato et al. 2013). Since all of the evaluated emergent constraints have been derived on the CMIP3 or CMIP5 ensemble, out-of-sample tests on the emergent constraints can be performed by assessing whether they still hold for the CMIP6 models. In this study, which is shown in **??** and already published in Schlund et al. (2020b), a substantial reduction of skill for the majority of emergent constraints is found when applied to the CMIP6 ensemble in comparison to the CMIP5 ensemble. This drop in skill is expressed as a decrease of the coefficient of determination R^2 of the emergent relationship and a decrease of the statistical significance using the null hypothesis that there is no correlation between the predictor and ECS. Moreover, the corresponding probability density functions (PDFs) for the emergent constraints show higher 66 % ECS ranges (17–83 % confidence) for almost all emergent constraints, resulting in values of 1.32–2.70 K for CMIP6 (CMIP5: 1.16–1.75 K). Averaged over all emergent constraints, this is an increase in the 66 % ECS range of 51 %. Similarly, the best estimates for ECS show values of 2.97–3.88 K in CMIP5 and 3.48–4.32 K in CMIP6, resulting in an increase of about 12 % averaged over all emergent constraints. Thus, key science question 2 (“*Can uncertainties in climate sensitivity be reduced with observations using the emergent constraint approach?*”) needs to be answered with a “not very well” for the CMIP6 ensemble. The increased best estimates and spreads resulting from the emergent constraints

in CMIP6 are likely related to the increased MMM and multi-model spread of ECS in CMIP6. A possible reason for the reduced skill of the emergent constraints when applied to the CMIP6 ensemble is the increased complexity of the CMIP6 models: A basic assumption for these single-process-oriented emergent constraints is that a single observable process dominates the uncertainty in ECS, which might not be valid anymore due to an increased number of processes that are included in the CMIP6 models.

To overcome these issues of single-process-oriented emergent constraints, **??** introduces an alternative approach based on ML. This work is already published in Schlund et al. (2020a). Since the new technique relies on a large number of data points in order to train the ML algorithm, the scalar climate sensitivity expressed as ECS or TCR is not an appropriate target variable. Therefore, this analysis does not focus on reducing uncertainties in climate sensitivity itself but rather on a selected process that contributes to it: gross primary production (GPP). GPP is the largest flux of the terrestrial carbon uptake and slows down global warming by removing CO_2 from the atmosphere. In the first step of the new two-step approach, the global mean GPP at the end of the 21st century in the CMIP5 RCP8.5 scenario is constrained to (171 ± 12) GtC yr^{-1} using a published emergent constraint by Wenzel et al. (2016a). This first step corrects the CMIP5 models' GPP response to CO_2, which shows a range of 156–247 GtC yr^{-1} in the raw multi-model ensemble. In the second step, a ML-based climate model weighting approach is used to further constrain the gridded GPP based on present-day predictors that are relevant for the simulation of GPP in the ESMs. The ML approach is mathematically similar to the multiple diagnostic ensemble regression (MDER) approach (Karpechko et al. 2013; Senftleben et al. 2020; Wenzel et al. 2016b), but additionally considers multi-dimensional (gridded) target variables and non-linear relationships between the predictors and the target variable. A a relationship between process-oriented predictors and future projections of GPP is established and then utilized to project today's observed conditions into the future. The prediction phase of the new method can be interpreted as an implicit performance weighting. However, due to the complex structure of the used ML algorithm (gradient boosted regression tree (GBRT)) it is not possible to extract specific values for the individual weights. Two target variables are considered: the gridded monthly climatologies of absolute GPP (2091–2100) and the gridded fractional GPP change over the 21st century (2100 versus 2000). The latter quantity shows an increased GPP change in the high latitudes compared to regions closer to the equator. The results of both approaches are consistent with each other and with the global constraint of the first step. The new approach is validated by comparing it to other statistical models (the CMIP5 MMM and a linear least absolute shrinkage and selection operator (LASSO) model) in a leave-one-model-out cross-validation (CV) setup. Compared to MMM (LASSO), a reduction of the resulting mean root mean square error of prediction (RMSEP) of up to 48 % (3 %) is found when using the ML approach. Moreover, the evaluation of the global and local feature importance allows further insights into the ML model. For the first target variable (absolute GPP), historical GPP is by far the most important predictor, which can be explained with a correction of the historical bias in GPP by the new approach. For the second variable (fractional change in GPP), near-surface air temperature (T)

and leaf area index (LAI) are the dominant features. This study directly addresses key science question 3 (“*Can uncertainties in multi-dimensional (gridded) climate projections be reduced with ML techniques and observations?*”), which can be confidently answered with “yes” based on the results found.

3.2 Outlook

Climate sensitivity is a policy-relevant and easy-to-use metric to assess the strength of climate change. The range of model results for this important climate metric, however, remains large and has not been narrowed over the last decades. In the latest generation of climate models contributing to CMIP6, the range has even increased (Meehl et al. 2020). Chapter 2 presents two possible reasons for this: changes in the aerosol-cloud interaction and changes in the shortwave cloud phase change feedback over the Southern Ocean. While these are likely explanations, there might be additional relevant aspects that have not been analyzed in detail yet. This could include additional feedback processes that are not present in older CMIP generations. Obvious candidates for these feedback processes are cloud-related feedbacks, which pose a major source of uncertainty in climate sensitivity in modern-day ESMs (Boucher et al. 2013). Identifying and quantifying such feedback mechanisms can further help to gain a better understanding of the CMIP6 models and potentially to reduce associated uncertainties in the entire multi-model ensemble. A potentially interesting process relevant in this context is the influence of the midlatitude jet position on clouds and CREs (Grise and Medeiros 2016).

A promising way forward to reduce uncertainties in multi-model climate projections is to apply the new flexible ML-based climate model weighting approach introduced in **??** to other target variables. While the scalar metrics ECS or TCR with one value per climate model do not offer enough training points for this technique, a possible alternative could be the gridded near-surface air temperature over the 21st century in the Shared Socioeconomic Pathway (SSP) scenarios. A relevant study in this context is given by Brunner et al. (2020), who use performance and interdependence-based climate model weighting to constrain the near-surface air temperature over the 21st century based on observable process-based predictors of today’s climate and which could serve as a possible baseline for the new ML-based weighting approach. An exciting and valuable addition to the method could be causal inference (Nowack et al. 2020; Runge et al. 2019). In the current approach, the relationships between the different predictors and the target variable are based on statistical correlation alone. Moreover, the local feature importance maps do not reveal true causal relationships between the predictors and the target variable, but only show the relative weight that is given to a specific predictor at a specific location for the prediction of the target variable by the ML model. By integrating causal networks into the new approach, it might be possible to explore and utilize true causal connections instead.

The massive progress in artificial intelligence (AI) and ML in recent years in combination with ever increasing computational power and resources provided by modern supercomputers has the potential for enormous improvements in climate modeling and analysis (Reichstein

et al. 2019). Apart from the presented ML-based weighting approach, further possible applications of AI and ML in climate science involve for example ML-based parameterizations that are learned from high-resolution climate models (Gentine et al. 2018; Rasp et al. 2018), ML-based analysis and prediction of forcing patterns (Barnes et al. 2019; Mansfield et al. 2020) or learning of entire ESMs from observational products (Geer 2021). Therefore, the foundation for innovative and groundbreaking research in climate science for the near future has been laid, which will help to further understand the Earth system and fight one of the greatest challenges for humankind today: climate change.

List of Abbreviations

AI artificial intelligence 20

AR Assessment Report 6

CESM Community Earth System Model

CH_4 methane 1

CMIP Coupled Model Intercomparison Project 5

CO_2 carbon dioxide 6

CRE cloud radiative effect 13

CV cross-validation 19

ECS effective climate sensitivity 5

ESGF Earth System Grid Federation 5

ESM Earth system model 5

ESMValTool Earth System Model Evaluation Tool 5

GBRT gradient boosted regression tree 19

GHG greenhouse gas 11

GPP gross primary production 19

GSAT global mean near-surface air temperature 5

HadGEM Hadley Centre Global Environmental Model 11

IPCC Intergovernmental Panel on Climate Change 5

KUM Cape Kumukahi, Hawaii

LAI leaf area index 20

LASSO least absolute shrinkage and selection operator 19

MDER multiple diagnostic ensemble regression 19

ML machine learning 17

List of Figures

List of Tables

References

Albritton, D. L., Filho, L. G. M., Cubasch, U., Dai, X., Ding, Y., Griggs, D. J., Hewitson, B., Houghton, J. T., Isaksen, I., Karl, T., McFarland, M., Meleshko, V. P., Mitchell, J. F. B., Noguer, M., Nyenzi, B. S., Oppenheimer, M., Penner, J. E., Pollonais, S., Stocker, T., & Trenberth, K. E. (2001). Technical Summary. Cambridge University Press. https://archive.ipcc.ch/ipccreports/tar/wg1/pdf/WG1_TAR-FRONT.PDF

Allen, M. R., & Ingram, W. J. (2002). Constraints on future changes in climate and the hydrologic cycle. *Nature*, *419*(6903), 224–232. https://doi.org/10.1038/nature01092

Andrews, T., Andrews, M. B., Bodas-Salcedo, A., Jones, G. S., Kuhlbrodt, T., Manners, J., Menary, M. B., Ridley, J., Ringer, M. A., Sellar, A. A., Senior, C. A., & Tang, Y. (2019). Forcings, Feedbacks, and Climate Sensitivity in HadGEM3-GC3.1 and UKESM1. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4377–4394. https://doi.org/10.1029/2019ms001866

Barnes, E. A., Hurrell, J. W., Ebert-Uphoff, I., Anderson, C., & Anderson, D. (2019). Viewing Forced Climate Patterns Through an AI Lens. *Geophysical Research Letters*, *46*(22), 13389–13398. https://doi.org/10.1029/2019gl084944

Bindoff, N. L., Stott, P. A., AchutaRao, K. M., Allen, M. R., Gillett, N., Gutzler, D., Hansingo, K., Hegerl, G., Hu, Y., Jain, S., Mokhov, I. I., Overland, J., Perlwitz, J., Sebbari, R., & Zhang, X. (2013). Detection and Attribution of Climate Change: from Global to Regional. Cambridge University Press. https://www.ipcc.ch/site/assets/uploads/2018/02/WG1AR5_Chapter10_FINAL.pdf

Bock, L., Lauer, A., **Schlund**, **M.**, Barreiro, M., Bellouin, N., Jones, C., Meehl, G. A., Predoi, V., Roberts, M. J., & Eyring, V. (2020). Quantifying Progress Across Different CMIP Phases With the ESMValTool. *Journal of Geophysical Research: Atmospheres*, *125*(21). https://doi.org/10.1029/2019jd032321

Bodas-Salcedo, A., Hill, P. G., Furtado, K., Williams, K. D., Field, P. R., Manners, J. C., Hyder, P., & Kato, S. (2016). Large Contribution of Supercooled Liquid Clouds to the Solar Radiation Budget of the Southern Ocean. *Journal of Climate*, *29*(11), 4213–4228. https://doi.org/10.1175/jcli-d-15-0564.1

Bodas-Salcedo, A., Mulcahy, J. P., Andrews, T., Williams, K. D., Ringer, M. A., Field, P. R., & Elsaesser, G. S. (2019). Strong Dependence of Atmospheric Feedbacks on Mixed-Phase Microphysics and Aerosol-Cloud Interactions in HadGEM3. *Journal of Advances in Modeling Earth Systems*, *11*(6), 1735–1758. https://doi.org/10.1029/2019ms001688

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., & Lohmann, U. (2013). Clouds and Aerosols. Cambridge University

Press. https://www.ipcc.ch/site/assets/uploads/2018/02/WG1AR5_Chapter07_FINAL-1.pdf

Bretherton, F. P., Bryan, K., & Woods, J. D. (1990). Time-Dependent Greenhouse-Gas-Induced Climate Change. Cambridge University Press. https://archive.ipcc.ch/ipccreports/far/wg_I/ipcc_far_wg_I_chapter_06.pdf

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

Charney, J. G., Arakawa, A., Baker, D. J., Bolin, B., Dickinson, R. E., Goody, R. M., Leith, C. E., Stommel, H. M., & Wunsch, C. I. (1979). *Carbon dioxide and climate: a scientific assessment*.

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W. J., Johns, T., & Krinner, G. (2013). Long-term Climate Change: Projections, Commitments and Irreversibility. Cambridge University Press. https://www.ipcc.ch/site/assets/uploads/2018/02/WG1AR5_Chapter12_FINAL.pdf

Cubasch, U., Meehl, G. A., Boer, G. J., Stouffer, R. J., Dix, M., Noda, A., Senior, C. A., Raper, S., & Yap, K. S. (2001). Projections of Future Climate Change. Cambridge University Press. https://archive.ipcc.ch/ipccreports/tar/wg1/pdf/TAR-09.PDF

Dittus, A. J., Hawkins, E., Wilcox, L. J., Sutton, R. T., Smith, C. J., Andrews, M. B., & Forster, P. M. (2020). Sensitivity of Historical Climate Simulations to Uncertain Aerosol Forcing. *Geophysical Research Letters*, *47*(13). https://doi.org/10.1029/2019gl085806

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

Eyring, V., Bock, L., Lauer, A., Righi, M., **Schlund**, **M.**, Andela, B., Arnone, E., Bellprat, O., Brötz, B., Caron, L.-P., Carvalhais, N., Cionni, I., Cortesi, N., Crezee, B., Davin, E. L., Davini, P., Debeire, K., de Mora, L., Deser, C., . . . Zimmermann, K. (2020). Earth System Model Evaluation Tool (ESMValTool) v2.0 – an extended set of large-scale diagnostics for quasi-operational and comprehensive evaluation of Earth system models in CMIP. *Geoscientific Model Development*, *13*(7), 3383–3438. https://doi.org/10.5194/gmd-13-3383-2020

Flato, G. M., 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., & Rummukainen, M. (2013). Evaluation of Climate Models. Cambridge University Press. https://www.ipcc.ch/site/assets/uploads/2018/02/WG1AR5_Chapter09_FINAL.pdf

Forster, P. M., Andrews, T., Good, P., Gregory, J. M., Jackson, L. S., & Zelinka, M. (2013). Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models. *Journal of Geophysical Research: Atmospheres*, *118*(3), 1139–1150. https://doi.org/10.1002/jgrd.50174

Forster, P. M., Maycock, A. C., McKenna, C. M., & Smith, C. J. (2020). Latest climate models confirm need for urgent mitigation. *Nature Climate Change*, *10*(1), 7–10. https://doi.org/10.1038/s41558-019-0660-0

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., ... Zaehle, S. (2020). Global Carbon Budget 2020. *Earth System Science Data*, *12*(4), 3269–3340. https://doi.org/10.5194/essd-12-3269-2020

Furtado, K., Field, P. R., Boutle, I. A., Morcrette, C. J., & Wilkinson, J. M. (2016). A Physically Based Subgrid Parameterization for the Production and Maintenance of Mixed-Phase Clouds in a General Circulation Model. *Journal of the Atmospheric Sciences*, *73*(1), 279–291. https://doi.org/10.1175/jas-d-15-0021.1

Geer, A. J. (2021). Learning earth system models from observations: machine learning or data assimilation? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *379*(2194), 20200089. https://doi.org/10.1098/rsta.2020.0089

Gentine, P., Pritchard, M., Rasp, S., Reinaudi, G., & Yacalis, G. (2018). Could Machine Learning Break the Convection Parameterization Deadlock? *Geophysical Research Letters*, *45*(11), 5742–5751. https://doi.org/10.1029/2018gl078202

Gettelman, A., Hannay, C., Bacmeister, J. T., Neale, R. B., Pendergrass, A. G., Danabasoglu, G., Lamarque, J. F., Fasullo, J. T., Bailey, D. A., Lawrence, D. M., & Mills, M. J. (2019). High Climate Sensitivity in the Community Earth System Model Version 2 (CESM2). *Geophysical Research Letters*, *46*(14), 8329–8337. https://doi.org/10.1029/2019gl083978

Gregory, J. M., Ingram, W. J., Palmer, M. A., Jones, G. S., Stott, P. A., Thorpe, R. B., Lowe, J. A., Johns, T. C., & Williams, K. D. (2004). A new method for diagnosing radiative forcing and climate sensitivity. *Geophysical Research Letters*, *31*(3), L03205. https://doi.org/10.1029/2003gl018747

Grise, K. M., & Medeiros, B. (2016). Understanding the Varied Influence of Midlatitude Jet Position on Clouds and Cloud Radiative Effects in Observations and Global Climate Models. *Journal of Climate*, *29*(24), 9005–9025. https://doi.org/10.1175/jcli-d-16-0295.1

Hamilton, D. S., Lee, L. A., Pringle, K. J., Reddington, C. L., Spracklen, D. V., & Carslaw, K. S. (2014). Occurrence of pristine aerosol environments on a polluted planet. *Proceedings of the National Academy of Sciences*, *111*(52), 18466–18471. https://doi.org/10.1073/pnas.1415440111

Haustein, K., Allen, M. R., Forster, P. M., Otto, F. E. L., Mitchell, D. M., Matthews, H. D., & Frame, D. J. (2017). A real-time Global Warming Index. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-14828-5

Hoose, C., Kristjánsson, J. E., Chen, J.-P., & Hazra, A. (2010). A Classical-Theory-Based Parameterization of Heterogeneous Ice Nucleation by Mineral Dust, Soot, and Biological Particles in a Global Climate Model. *Journal of the Atmospheric Sciences*, *67*(8), 2483–2503. https://doi.org/10.1175/2010jas3425.1

Hyder, P., Edwards, J. M., Allan, R. P., Hewitt, H. T., Bracegirdle, T. J., Gregory, J. M., Wood, R. A., Meijers, A. J. S., Mulcahy, J., Field, P., Furtado, K., Bodas-Salcedo, A., Williams, K. D., Copsey, D., Josey, S. A., Liu, C., Roberts, C. D., Sanchez, C., Ridley, J., . . . Belcher, S. E. (2018). Critical Southern Ocean climate model biases traced to atmospheric model cloud errors. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-05634-2

IPCC. (2013). Summary for Policymakers. Cambridge University Press. https://www.ipcc.ch/site/assets/uploads/2018/02/WG1AR5_SPM_FINAL.pdf

IPCC. (2014). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. IPCC. https://www.ipcc.ch/site/assets/uploads/2018/02/SYR_AR5_FINAL_full.pdf

Karpechko, A. Y., Maraun, D., & Eyring, V. (2013). Improving Antarctic Total Ozone Projections by a Process-Oriented Multiple Diagnostic Ensemble Regression. *Journal of the Atmospheric Sciences*, *70*(12), 3959–3976. https://doi.org/10.1175/Jas-D-13-071.1

Kattenberg, A., Giorgi, F., Grassl, H., Meehl, G. A., Mitchell, J. F. B., Stouffer, R. J., Tokioka, T., Weaver, A. J., & Wigley, T. M. L. (1996). Climate models – projections of future climate. Cambridge University Press. https://archive.ipcc.ch/ipccreports/sar/wg_I/ipcc_sar_wg_I_full_report.pdf

Kay, J. E., Bourdages, L., Miller, N. B., Morrison, A., Yettella, V., Chepfer, H., & Eaton, B. (2016). Evaluating and improving cloud phase in the Community Atmosphere Model version 5 using spaceborne lidar observations. *Journal of Geophysical Research: Atmospheres*, *121*(8), 4162–4176. https://doi.org/10.1002/2015jd024699

Kiehl, J. T. (2007). Twentieth century climate model response and climate sensitivity. *Geophysical Research Letters*, *34*(22). https://doi.org/10.1029/2007gl031383

Knutti, R., Sedlacek, J., Sanderson, B. M., Lorenz, R., Fischer, E. M., & Eyring, V. (2017). A climate model projection weighting scheme accounting for performance and interdependence. *Geophysical Research Letters*, *44*(4), 1909–1918. https://doi.org/10.1002/2016gl072012

Lauer, A., Eyring, V., Bellprat, O., Bock, L., Gier, B. K., Hunter, A., Lorenz, R., Pérez-Zanón, N., Righi, M., **Schlund**, **M.**, Senftleben, D., Weigel, K., & Zechlau, S. (2020). Earth System Model Evaluation Tool (ESMValTool) v2.0 – diagnostics for emergent constraints and future projections from Earth system models in CMIP. *Geoscientific Model Development*, *13*(9), 4205–4228. https://doi.org/10.5194/gmd-13-4205-2020

Mansfield, L. A., Nowack, P. J., Kasoar, M., Everitt, R. G., Collins, W. J., & Voulgarakis, A. (2020). Predicting global patterns of long-term climate change from short-term simulations using machine learning. *npj Climate and Atmospheric Science*, *3*(1). https://doi.org/10.1038/s41612-020-00148-5

McCoy, D. T., Hartmann, D. L., Zelinka, M. D., Ceppi, P., & Grosvenor, D. P. (2015). Mixed-phase cloud physics and Southern Ocean cloud feedback in climate models. *Journal of Geophysical Research: Atmospheres*, *120*(18), 9539–9554. https://doi.org/10.1002/2015jd023603

McCoy, D. T., Tan, I., Hartmann, D. L., Zelinka, M. D., & Storelvmo, T. (2016). On the relationships among cloud cover, mixed-phase partitioning, and planetary albedo in GCMs. *Journal of Advances in Modeling Earth Systems*, *8*(2), 650–668. https://doi.org/10.1002/2015ms000589

Meehl, G. A., Senior, C. A., Eyring, V., Flato, G., Lamarque, J.-F., Stouffer, R. J., Taylor, K. E., & **Schlund**, **M.** (2020). Context for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models. *Science Advances*, *6*(26), eaba1981. https://doi.org/10.1126/sciadv.aba1981

Meyers, M. P., DeMott, P. J., & Cotton, W. R. (1992). New Primary Ice-Nucleation Parameterizations in an Explicit Cloud Model. *Journal of Applied Meteorology*, *31*(7), 708–721. https://doi.org/10.1175/1520-0450(1992)031<0708:npinpi>2.0.co;2

Mitchell, J. F. B., Manabe, S., Meleshko, V., & Tokioka, T. (1990). Equilibrium Climate Change and its Implications for the Future. https://archive.ipcc.ch/ipccreports/far/wg_I/ipcc_far_wg_I_chapter_05.pdf

Nowack, P., Runge, J., Eyring, V., & Haigh, J. D. (2020). Causal networks for climate model evaluation and constrained projections. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-15195-y

Randall, D. A., Wood, R. A., Bony, S., Colman, R., Fichefet, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., Stouffer, R. J., Sumi, A., & Taylor, K. E. (2007). Climate Models and Their Evaluation. Cambridge University Press. https://www.ipcc.ch/site/assets/uploads/2018/02/ar4-wg1-chapter8-1.pdf

Rasp, S., Pritchard, M. S., & Gentine, P. (2018). Deep learning to represent subgrid processes in climate models. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(39), 9684–9689. https://doi.org/10.1073/pnas.1810286115

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, *566*(7743), 195–204. https://doi.org/10.1038/s41586-019-0912-1

Righi, M., Andela, B., Eyring, V., Lauer, A., Predoi, V., **Schlund**, **M.**, Vegas-Regidor, J., Bock, L., Brotz, B., de Mora, L., Diblen, F., Dreyer, L., Drost, N., Earnshaw, P., Hassler, B., Koldunov, N., Little, B., Tomas, S. L., & Zimmermann, K. (2020). Earth System Model Evaluation Tool (ESMValTool) v2.0-technical overview. *Geoscientific Model Development*, *13*(3), 1179–1199. https://doi.org/10.5194/gmd-13-1179-2020

Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., & Sejdinovic, D. (2019). Detecting and quantifying causal associations in large nonlinear time series datasets. *Science Advances*, *5*(11), https://advances.sciencemag.org/content/5/11/eaau4996.full.pdf. https://doi.org/10.1126/sciadv.aau4996

Schlund, **M.**, Eyring, V., Camps-Valls, G., Friedlingstein, P., Gentine, P., & Reichstein, M. (2020a). Constraining Uncertainty in Projected Gross Primary Production With Machine Learning. *Journal of Geophysical Research: Biogeosciences*, *125*(11), e2019JG005619. https://doi.org/10.1029/2019jg005619

Schlund, M., Lauer, A., Gentine, P., Sherwood, S. C., & Eyring, V. (2020b). Emergent constraints on equilibrium climate sensitivity in CMIP5: do they hold for CMIP6? *Earth System Dynamics*, *11*(4), 1233–1258. https://doi.org/10.5194/esd-11-1233-2020

Senftleben, D., Lauer, A., & Karpechko, A. (2020). Constraining Uncertainties in CMIP5 Projections of September Arctic Sea Ice Extent with Observations. *Journal of Climate*, *33*(4), 1487–1503. https://doi.org/10.1175/jcli-d-19-0075.1

Solomon, S., Qin, D., Manning, M., Alley, R. B., Berntsen, T., Bindoff, N. L., Chen, Z., Chidthaisong, A., Gregory, J. M., Hegerl, G. C., Heimann, M., Hewitson, B., Hoskins, B. J., Joos, F., Jouzel, J., Kattsov, V., Lohmann, U., Matsuno, T., Molina, M., ... Wratt, D. (2007). Technical Summary. Cambridge University Press. https://www.ipcc.ch/site/assets/uploads/2018/02/ar4-wg1-ts-1.pdf

Stocker, T. F., Qin, D., Plattner, G.-K., Alexander, L. V., Allen, S. K., Bindoff, N. L., Bréon, F.-M., Church, J. A., Cubasch, U., Emori, S., Forster, P., Friedlingstein, P., Gillett, N., Gregory, J. M., Hartmann, D. L., Jansen, E., Kirtman, B., Knutti, R., Kumar, K. K., ... Xie, S.-P. (2013). Technical Summary. Cambridge University Press. https://www.ipcc.ch/site/assets/uploads/2018/02/WG1AR5_TS_FINAL.pdf

Tan, I., Storelvmo, T., & Zelinka, M. D. (2016). Observational constraints on mixed-phase clouds imply higher climate sensitivity. *Science*, *352*(6282), 224–227. https://doi.org/10.1126/science.aad5300

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An Overview of Cmip5 and the Experiment Design. *Bulletin of the American Meteorological Society*, *93*(4), 485–498. https://doi.org/10.1175/Bams-D-11-00094.1

UNFCCC. (2015). Adoption of the Paris Agreement. http://unfccc.int/resource/docs/2015/cop21/eng/l09r01.pdf

WCRP. (2020). *WCRP Coupled Model Intercomparison Project (CMIP)*. https://www.wcrp-climate.org/wgcm-cmip

Weigel, K., Bock, L., Gier, B. K., Lauer, A., Righi, M., **Schlund, M.**, Adeniyi, K., Andela, B., Arnone, E., Berg, P., Caron, L.-P., Cionni, I., Corti, S., Drost, N., Hunter, A., Lledó, L., Mohr, C. W., Paçal, A., Pérez-Zanón, N., ... Eyring, V. (2020). Earth System Model Evaluation Tool (ESMValTool) v2.0 – diagnostics for extreme events, regional and impact evaluation and analysis of Earth system models in CMIP. *Geoscientific Model Development Discussions*, *in review*, 1–43. https://doi.org/10.5194/gmd-2020-244

Wenzel, S., Cox, P. M., Eyring, V., & Friedlingstein, P. (2016a). Projected land photosynthesis constrained by changes in the seasonal cycle of atmospheric CO2. *Nature*, *538*(7626), 499–501. https://doi.org/10.1038/nature19772

Wenzel, S., Eyring, V., Gerber, E. P., & Karpechko, A. Y. (2016b). Constraining Future Summer Austral Jet Stream Positions in the CMIP5 Ensemble by Process-Oriented Multiple Diagnostic Regression. *Journal of Climate*, *29*(2), 673–687. https://doi.org/10.1175/Jcli-D-15-0412.1

Williams, K. D., Hewitt, A. J., & Bodas-Salcedo, A. (2020). Use of Short-Range Forecasts to Evaluate Fast Physics Processes Relevant for Climate Sensitivity. *Journal of Advances in Modeling Earth Systems*, *12*(4). https://doi.org/10.1029/2019ms001986

Wyser, K., van Noije, T., Yang, S. T., von Hardenberg, J., O'Donnell, D., & Doscher, R. (2020). On the increased climate sensitivity in the EC-Earth model from CMIP5 to CMIP6. *Geoscientific Model Development*, *13*(8), 3465–3474. https://doi.org/10.5194/gmd-13-3465-2020

Zelinka, M. D., Myers, T. A., Mccoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., Klein, S. A., & Taylor, K. E. (2020). Causes of Higher Climate Sensitivity in CMIP6 Models. *Geophysical Research Letters*, *47*(1), e2019GL085782. https://doi.org/10.1029/2019GL085782