https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Fresh Eyes on CMIP Model Biases: Diagnosing Biases Across CMIP Generations and Implications for CMIP7

Flossie Brown[1], Sabine Bischof[2], Brodie Pearson[3], Aditya Sengupta[4], Pauline Bonnet[5], Swen Brands[6], Jayasankar C.B.[7], James Carruthers[8], Marion Devilliers[9], F. Feba[10], Mehmet Sedat Gözlet[2], Saurabh Kelkar[11], Sifan A. Koriche[12], Michael Lai[13], Andre Lanfer Marquez[14], Xiaohan Sally Li[15], Yiwen Li[16], Abhnil Prasad[17], Angelica R. Rodriguez[18], Serena Schroeter[19], Gopika Suresh[20], and Rebecca Varney[21]

[1]Institute for Atmospheric and Climate Science, ETH Zurich, 8092 Zurich, Switzerland
[2]GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany
[3]College of Earth, Ocean, & Atmospheric Sciences, Oregon State University, Corvallis, OR, United States of America
[4]School of Geography, Earth and Atmospheric Sciences, University of Melbourne, Melbourne, Kulin nations, Australia
[5]Seismology team of IPGP, Centre des Matériaux of Mines ParisTech, PIMM laboratory in ENSAM.
[6]Instituto de Física de Cantabria (CSIC-UC), Santander, Spain
[7]College of Marine Science, University of South Florida, St. Petersburg, FL, USA
[8]School of Engineering, Newcastle University, Newcastle upon Tyne, United Kingdom
[9]Danish Meteorological Institute, Copenhagen, Denmark
[10]Laboratoire de Météorologie Dynamique (LMD), École Polytechnique, Palaiseau, France
[11]WOTR Centre for Resilience Studies, Watershed Organisation Trust (WOTR), Pune, India
[12]Alabama Water Institute, University of Alabama, Tuscaloosa, USA
[13]Met Office Hadley Centre, Exeter, UK
[14]National Institute for Space Research - INPE, São José dos Campos, Brazil
[15]Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, NJ, USA
[16]China University of Geosciences (Beijing), Beijing
[17]Climate Change Research Centre, University of New South Wales, Sydney, New South Wales, Australia
[18]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California
[19]CSIRO Environment, Hobart, Tas., Australia
[20]National Center For Medium Range Weather Forecasting, Noida, India
[21]Department of Physical Geography and Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden

Correspondence: Sabine Bischof (sbischof@geomar.de)

Abstract. The Coupled Model Intercomparison Project (CMIP) simulates standardized experiments spanning historical, future, and hypothetical conditions, to better understand the Earth system's evolution. Systematic biases remain a persistent limitation of coupled climate models. Within the framework of the CMIP, successive model generations have increased in complexity, resolution, and representation of Earth system processes, yet many long-standing biases remain across atmosphere, ocean, land, and cryosphere components, and their interactions and feedbacks. In this literature review, we identify the characteristics of key systematic biases in coupled models from CMIP6 and earlier CMIP phases to aid future comparisons to CMIP7. We introduce results of a community survey, designed to prioritize diagnostics for the Rapid Evaluation Framework (REF), and use these diagnostics for our review. Biases are shown to be interconnected across the Earth system, via sea surface temperature distributions, ENSO patterns, AMOC´s large scale transport biases, sea ice, the carbon cycle and radiation fluxes among others,

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

with impacts on the Earth system response through metrics such as the Transient Climate Response and Equilibrium Climate sensitivity. The physical mechanisms underlying these biases and strategies to reduce them are presented, including improved parameterizations, increased resolution, and enhanced coupling among Earth system components. This review synthesizes current understandings of systematic biases across the Earth System, providing an assessment of biases in a set of critical diagnostics. In addition we provide a perspective on the advances in physical processes, higher resolution modelling frameworks, and targeted experiments and model evaluation frameworks expected in CMIP7. This review serves as an overview for end users, those new to CMIP data and those looking to connect different aspects of the Earth system.

1 Introduction

The Coupled Model Intercomparison Project (CMIP) has existed for almost 30 years and has evolved through 5 major phases, with another phase planned in alignment with the next report from the Intergovernmental Panel on Climate Change (IPCC). In the most recent iteration, CMIP6, dozens of unique coupled climate models, encompassing General Circulation Models (GCMs) as well as Earth System Models (ESMs) simulated standardized experiments spanning historical, future, and experimental conditions, to better understand the Earth system's evolution. Crucially, CMIP's future scenario experiments allow us to project the range of conditions that Earth may experience as a function of different climate forcings. To provide context for these projections, it is important to understand how CMIP models differ from the real Earth system.

Through all the phases of CMIP there have been systematic model biases, where the behavior of historical experiments deviate from observations of the Earth system. Throughout the different CMIP phases, participating coupled climate models have become more complex, with finer resolution, higher extension into the atmosphere, an increased number of coupled Earth system components, and more advanced parameterizations for unresolved physical, chemical, biological and geological processes. These advances have reduced many systematic model biases, but there are still many biases that persist. For example, Stouffer et al. (2017) highlighted several long-standing biases spanning the ocean, atmosphere, land and their coupling, that were still present in CMIP6. Ongoing model development is expected to reduce biases further for example by increased model resolution. However, increasing complexity, such as including an active carbon cycle or ice-sheets might introduce new uncertainties. The focus of CMIP7 on emission-driven rather than concentration-driven future scenarios will increase the effect of how well the carbon cycle is represented in coupled climate models.

One aim of the CMIP6 experiments was to identify "What are the origins and consequences of systematic model biases?" (Eyring et al., 2016). Similarly, in anticipation of CMIP7, the Rapid Evaluation Framework (REF) was developed by the Model Benchmarking and Evaluation Task Team to efficiently evaluate and identify biases in key Earth System diagnostics (Hoffman et al., 2025). The REF platform will enable comparison between models and across idealized and historical experiments against observational evidences. The first set of fast track REF diagnostics span the five themes of the CMIP7 Data Request: Oceans and Sea Ice, Land and Land Ice, Atmosphere, Earth System, and Impacts and Adaptation (Hoffman et al., 2025). As part of the diagnostic selection process, a community-wide survey was conducted. The results of this survey are presented here, and used as a basis for describing model biases across the CMIP7 themes.

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Here we present the community survey results used to select REF diagnostics, and synthesize literature from CMIP6 and earlier CMIP experiments to explore the origins and consequences of systematic model biases across all components of the Earth system, including several of the longstanding biases identified by Stouffer et al. (2017). Where possible, we also quantify how these biases have changed between CMIP6 and earlier CMIP phases, and potential solutions to reduce these biases in CMIP7 and other future experiments.

2 Selection of diagnostics

2.1 Survey methods

The diagnostics selected to review here, align with the priority diagnostics selected for evaluation as part of the REF (Hoffman et al., 2025). A CMIP community-wide survey was conducted to identify these key diagnostics across five thematic areas: Atmosphere, Earth System, Land and Ice, Ocean and Sea Ice, and Impacts and Adaptations. A total of 47 individuals responded to the survey, the results of which are analysed here to guide diagnostic priorities for the REF in preparation for CMIP7.

Respondents were asked to rank up to five diagnostics within each theme from a list of suggested diagnostics developed by the CMIP Model Benchmarking and Evaluation Task Team (Appendix A). A rank of 1 indicates that the respondent considers it to be the most important diagnostic to evaluate first. The respondents also selected their preferred metrics to view the evaluation results. The metrics offered for comparison in the survey included 'bias', 'interannual variability', 'spatial distribution', 'seasonal cycle', and 'other periodic'. In this context, 'diagnostic' refers to a model output, or a quantity derived from it, that can be compared with observational data for the purposes of evaluation and benchmarking, whereas 'metric' describes the method of evaluation.

Beyond ranking diagnostics, participants were invited to provide general feedback on the REF, including suggestions for additional simulations, missing diagnostics, and observational reference datasets. A copy of the survey questions is provided in Appendix A.

2.1.1 Respondent Overview

The respondent pool reflected a broad cross-section of the climate modelling community. Approximately 40% identified as modelling centres, 30% as members of Model Intercomparison Projects (MIPs), 10% as part of data infrastructure, and 5% as observational data providers. Several respondents represented more than one category. Importantly, 16 of the 33 MIPs registered for CMIP7 were represented among the survey respondents (AerChemMIP, SIMIP, GeoMIP, DAMIP, ScenarioMIP, PMIP, CMIP DECK, VolMIP, DCPP, HighresMIP, NAHosMIP, DynVarMIP, PAMIP, C4MIP, MethaneMIP, and CFMIP).

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Table 1. The final diagnostics and themes chosen for priority evaluation in the REF. The column 'Section' gives the Section of this paper that the review can be found.

Section	Diagnostic	Theme	In original survey
1.1	Sea surface temperature (SST)/Sea surface salinity (SSS) bias	Oceans and sea ice	Yes
1.1	Ocean heat content (OHC)	Oceans and sea ice	Yes
1.2	Antarctic/Arctic sea ice area seasonal cycle	Oceans and sea ice	Yes
1.2	Antarctic annual mean/Arctic September rate of sea ice area (SIA) loss per degree warming (dSIA/dGMST)	Oceans and sea ice	Yes
1.3	Atlantic Meridional Overturning Circulation (AMOC)	Oceans and sea ice	Yes
1.4	El Nino Southern Oscillation (ENSO) diagnostics (lifecycle/ seasonality/ teleconnections)	Oceans and sea ice	Yes
2.1	Runoff	Land and land ice	Yes
2.1	Surface soil moisture (SM)	Land and land ice	Yes
2.2	Soil carbon	Land and land ice	Yes
2.2	Gross primary production (GPP)	Land and land ice	Yes
2.2	Net Ecosystem Exchange (NEE)	Land and land ice	Yes
2.2	Leaf area index (LAI)	Land and land ice	Yes
	Snow cover	Land and land ice	No
3.1	Radiative and heat fluxes at the surface and top of atmosphere (TOA)	Atmosphere	Yes
3.1	Cloud Radiative Effects	Atmosphere	No
3.1	Scatterplots of two cloud-relevant variables (specific regions of the globe: with specific cloud regimes)	Atmosphere	No
3.2	Annual cycle and seasonal mean of multiple variables	Atmosphere	Yes
3.2	Evaporation - Precipitation (E-P)	Atmosphere	Yes
3.2	Double Inter-tropical convergence zone (ITCZ)	Atmosphere	Yes
3.3	Climate variability modes (e.g., MJO, Extratropical modes of variability)	Atmosphere	Yes
4.1	Equilibrium climate sensitivity (ECS)	Earth System	Yes
4.1	Transient climate response (TCR)	Earth System	Yes
4.2	Transient climate response to cumulative emissions of carbon dioxide (TCRE)	Earth System	Yes
4.2	Zero Emissions Commitment (ZEC)	Earth System	Yes
	Historical changes in climate variables (time series, trends)	Earth System	Yes
5	High Amplitude Rossby Waves	Impacts and Adaptation	Yes
5	Climate drivers for fire (Fire burnt area, fire weather and fuel continuity)	Impacts and Adaptation	No
	Internal Variability or ensemble spread within individual models (focused on precipitation and surface temperature)	Impacts and Adaptation	Yes
	Evaluation of key climate variables at Global Warming Levels	Impacts and Adaptation	Yes

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

2.2 Survey results

2.2.1 Overall Feedback

Overall, the idea that the REF could be an efficient method for model evaluation was well received, with 75% of respondents rating it very positively (4 or 5 stars). However, some expressed concerns regarding the feasibility of certain diagnostics. In particular, diagnostics that require multiple variables, especially those outputs with differing units or formats across models, were seen as complex and potentially error-prone. Respondents emphasised a preference for focusing on simpler diagnostics that can be rapidly released, to better support timely evaluation and benchmarking.

2.2.2 Diagnostics and Ranking

Figure 1 shows the number of votes of each rank for the suggested diagnostics in the survey. Reflecting its broad impact on many aspects of the Earth system, 'Sea surface temperature (SST)/Sea surface salinity (SSS)' was considered the most important diagnostic of the Ocean and Sea Ice theme by number of Rank 1 votes, followed by 'Antarctic/Arctic sea ice area seasonal cycle'. These both received the highest total number of votes (Fig. 1).
'Gross primary production (GPP)' received the highest number of first-place and total rankings among the Land and Land Ice diagnostics. Surface soil moisture received a similar total number of votes, but only one respondent ranked it first. 'Surface mass balance (land ice)' was not considered highly important but respondents suggested 'snow cover' and 'vegetation biomass' as diagnostics they would also be interested in (Table 1).
In the Atmosphere theme, the diagnostics considered most important, both in terms of total votes and number of Rank 1 votes, were 'Annual cycle and seasonal means' and 'Radiative and heat fluxes' (Fig. 1). Additionally, three respondents ranked Evaporation-Precipitation as the most important diagnostic. Several respondents noted that important diagnostics related to aerosols and clouds were missing from the survey list.
Within the 'Earth System Diagnostics' theme, 'Historical changes in climate variables' was the most highly ranked diagnostic and 'Equilibrium climate sensitivity' received the second highest number of Rank 1 votes (Fig. 1). These diagnostics help us assess the model capabilities in simulating key Earth system processes (such as cloud formation, carbon cycle feedbacks, aerosol-cloud interactions etc.) which are often complex and a significant source of uncertainty in climate projections. Respondents also expressed an interest in weather extremes, particularly those related to precipitation and temperature.
For the 'impacts' theme, only three diagnostics were suggested. 'High amplitude Rossby waves' were not considered a high priority compared to the more general diagnostics 'Climate variables at global warming levels' and 'Internal variability and ensemble spread' that could be evaluated for a number of variables (Fig. 1).
Using suggestions of importance from the survey, a final list of priority diagnostics was prepared (Table 1) by the Model Benchmarking and Evaluation Task Team.

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

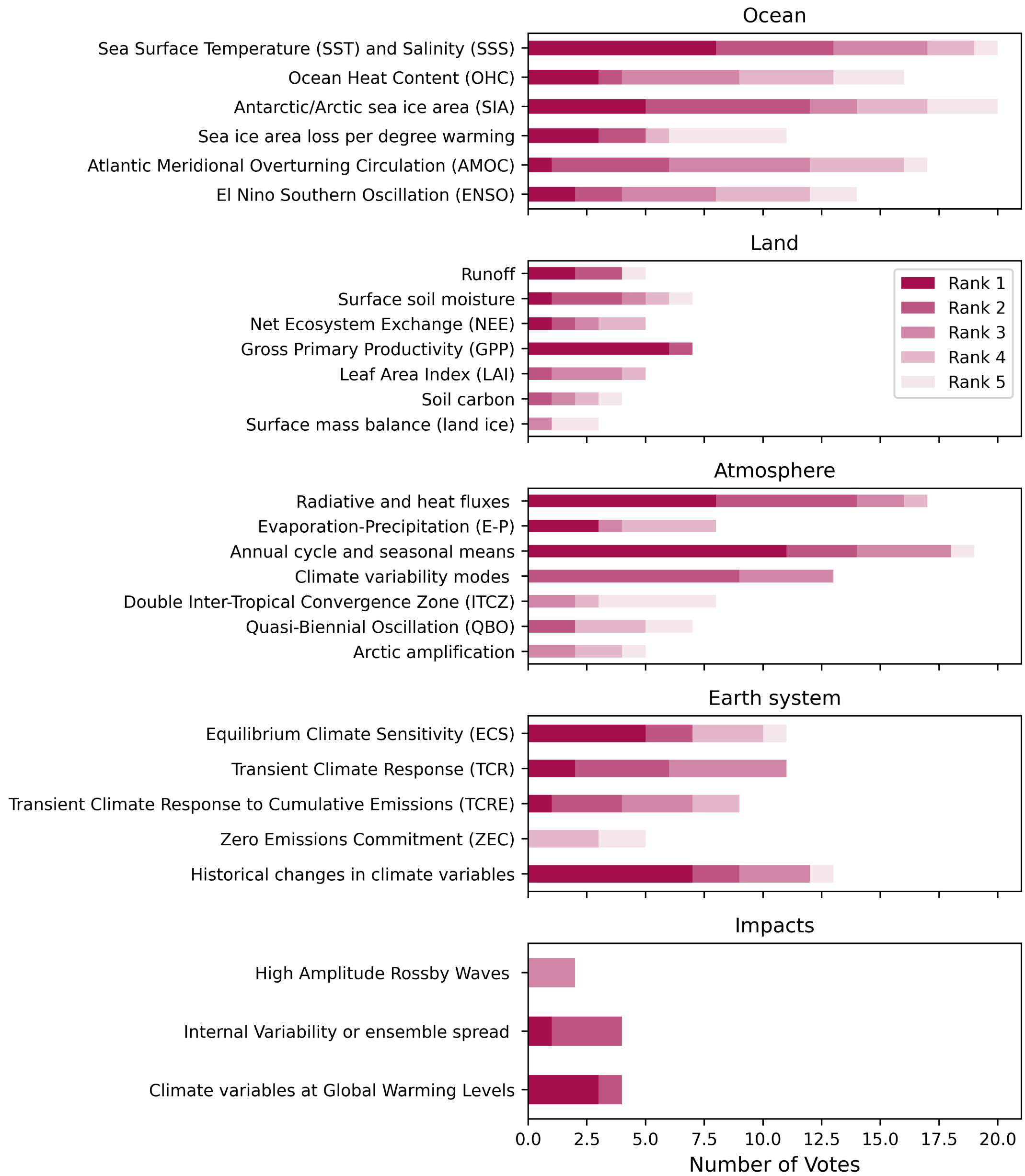

Figure 1. Number of votes for each proposed diagnostic and rank. Survey participants were asked to rank up to 5 diagnostics in each theme, with rank 1 indicating the most important diagnostic to evaluate first.

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

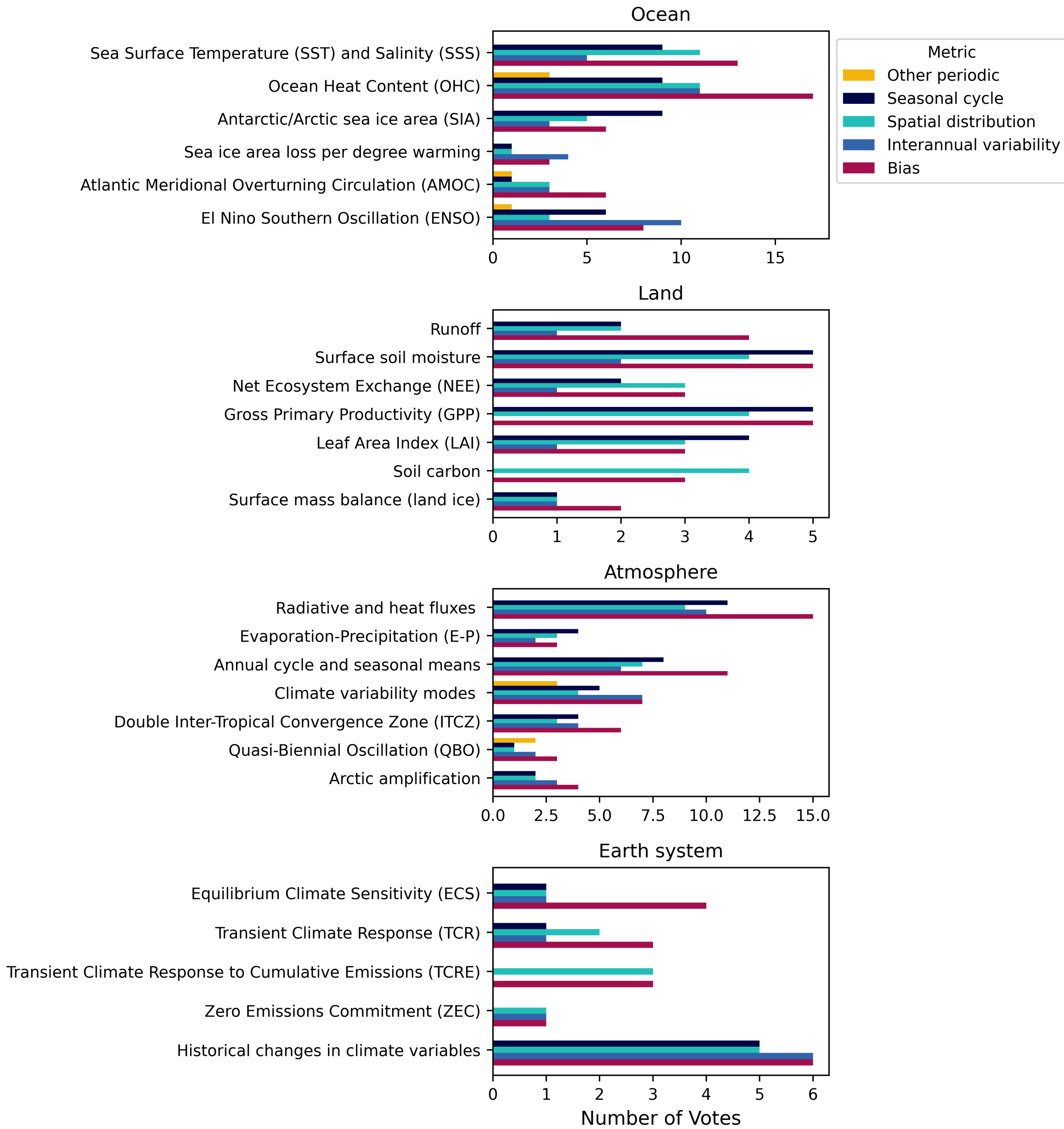

Figure 2. Number of votes for preferred metric of evaluation for each diagnostic. Survey participants were asked to indicate which evaluation metrics were important to calculate, with no limits on numbers of metrics they could select.

100 ### 2.2.3 Metrics

The survey results indicate that 'bias' was consistently the most preferred metric for presenting data across diagnostics (Fig. 2). There was interest in bias, interannual variability, spatial distributions and seasonal cycles for the Ocean and Sea Ice and

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Atmosphere variables, often with a slight overall leaning toward 'bias'. Being an annually changing phenomenon, ENSO had the highest number of votes for interannual variability. Similarly, 'other periodic' metrics were only recommended in cases where well-known periodicities were relevant, such as for the QBO and other climate modes. In contrast, Land and Land Ice diagnostics were voted to be most effectively presented as spatial distributions and through seasonal analysis. The Earth System diagnostics also received a high interest in spatial distribution.

2.2.4 Key findings from the survey and motivation to explore model biases

The results of the survey were used to inform the priority diagnostics to be included in the REF. Additional diagnostics suggested by survey respondents were also included (Table 1); the survey highlighted a demand for expanded cloud and aerosol diagnostics. In response, additional variables ('Cloud Radiative Effects' and 'Scatterplots of two cloud-relevant variables') have been incorporated into the list of diagnostics in this review (Table 1). A request for 'snow cover' was included in place of 'surface mass balance' and 'Climate drivers for fire' were added to the Impacts and Adaptations diagnostics. The highest ranking diagnostics are useful to understand which diagnostics are considered most important to a range of CMIP users, developers and modelling centres. However, the limited number of participants restricts further interpretation.
The strong preference for 'bias' as a metric for evaluation has prompted a literature review focused on the biases in key diagnostics identified here. Other suggested metrics, particularly those related to timescales (e.g., seasonal or interannual), are used to inform our presentation of the biases and their implications.

3 Biases in selected diagnostics

3.1 Ocean and Sea Ice

3.1.1 Ocean State

Biases in ocean temperature and salinity are important for a range of climatic reasons. Temperature and salinity variations affect the density of seawater, controlling the ocean's stratification and lateral pressure gradients. This impacts ocean overturning and circulation at a range of scales. Variations in the sea surface temperature and salinity mediate the fluxes of heat and freshwater at air-sea and ice-ocean interfaces (note that ocean salinity is related to freshwater content). These fluxes drive weather (i.e., hurricane formation) and climate (i.e., formation of sea ice; deep convection and thermohaline circulation).

Sea Surface Temperature

SST biases remain a persistent weakness of CMIP models, strongly influencing the coupled climate system. SST biases and the underlying process are region-dependent. CMIP models show warm biases in the southeast and northeast Pacific (Zhang et al., 2023), southeast Atlantic (Richter, 2015), and parts of the Southern Ocean (Lauer et al., 2018), while cold biases appear in the North Atlantic (e.g., Zhang et al., 2023), northwest Pacific (e.g., Bellenger et al., 2014), and Arabian Sea (e.g., Fathrio et al., 2017; McKenna et al., 2024). There are also CMIP6 biases in spatial SST gradients, for example in the tropical Indian Ocean

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

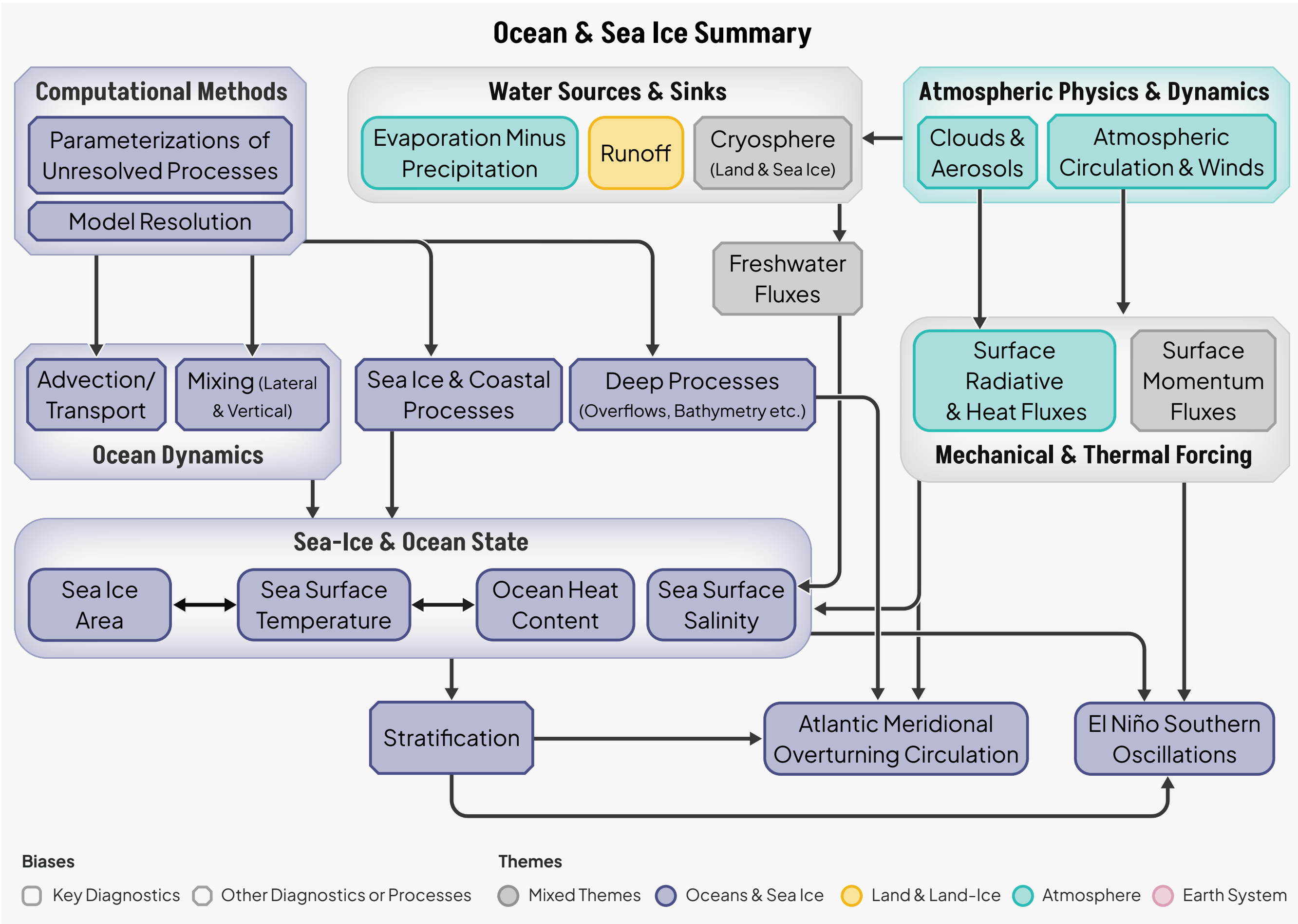

Figure 3. Key diagnostics and drivers of bias in the Ocean

during boreal winter (Zhang et al., 2026) or at inter-hemispheric scales (Zhang et al., 2024), and biases in spatial warming patterns (Armour et al., 2024).

Broadly, SST biases emerge from structural factors such as limited horizontal and vertical resolution (Fig. 3), deficiencies in subgrid-scale parameterizations, numerical diffusion, and inaccuracies in the representation of key oceanic and atmospheric processes, particularly in regions where observations are sparse and model constraints are weak especially in the deep ocean where observations are sparse (Huang et al., 2003; Hawkins and Sutton, 2009; Zheng et al., 2011; Exarchou et al., 2018; Zhang et al., 2023, 2024). These inconsistencies can introduce initial shocks and model drift that further contribute to SST errors. Regional SST biases arise from a combination of surface flux and dynamical errors, including surface winds and AMOC representation (see Sect. 3.1.3). Warm biases in the eastern tropical oceans result from insufficient vertical mixing, weak winds, and inaccurately represented cloud cover in CMIP models (Richter, 2015; Exarchou et al., 2018). The persisting Southern Ocean warm bias in several generations of CMIP models is mainly attributed to inaccurate representation of AMOC (Wang

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

et al., 2014), and clouds (Hyder et al., 2018). Similarly, the cold North Atlantic biases stem from a weak AMOC and reduced northward heat transport (Zhang and Zhao, 2015). The cold bias in the northwest Pacific is linked to excessive latent heat flux due to overestimated winds in CMIP (Wang et al., 2018). Surface and sub-surface temperature biases in the Indian Ocean are driven by biases in surface wind, local meridional overturning, and Indonesian Throughflow, and could be corrected by improved air-sea interactions and inter-basin transfers (Zhang et al., 2026; Gopika et al., 2025). The Arabian Sea cold bias is also attributed to overly strong surface winds (McKenna et al., 2024). SST biases directly affect the air-sea fluxes, precipitation distribution, circulation and the representation of key modes of climate variability (See Sect. 3.3.2, and 3.3.3). Biases in spatial SST patterns could limit the ability of models to reproduce the relation between global warming and climate sensitivity (Armour et al., 2024; Gopika et al., 2024).

Sea Surface Salinity

Sea surface salinity (SSS) in CMIP6 is generally too low (i.e., surface water is too fresh) at both basin and global scales (Liu et al., 2022; van Westen and Dijkstra, 2024). SSS biases can propagate to the ocean's density structure, influencing ocean dynamics and thermodynamics through altered horizontal and vertical temperature gradients (fronts and stratification respectively) (Fig. 3). Ocean density and related thermodynamics are especially sensitive to salinity variations in polar regions where water is cold, compared to more temperate regions where density fluctuations are largely driven by temperature changes. The global and Atlantic CMIP6 fresh biases have increased since CMIP5, attributed to a combination of deficient surface freshwater fluxes from atmospheric (P-E; see Section 3.3.2) and riverine sources, and to insufficient resolution or parameterization of lateral/isopycnal and vertical/diapycnal oceanic mixing (Liu et al., 2022). SSS biases could be alleviated through a combination of improved external freshwater fluxes, more accurate ocean mixing parameterizations, and higher-resolution simulations (Liu et al., 2022; van Westen and Dijkstra, 2024). In polar regions, the Arctic basin fresh bias was reduced between CMIP5 and CMIP6 but still persists (Wang et al., 2022b), and the Southern Ocean freshwater fluxes are underestimated in CMIP6 (Kaufman et al., 2025). These Arctic and Southern Ocean biases both result from precipitation and run-off errors, and could be alleviated by inclusion of Ice Sheets (Greenland and Antarctic respectively). The Indian Ocean is another region where SSS biases are important, with salinity stratification forming barrier layers and thin freshwater lenses near the ocean surface, mediating air-sea fluxes, and upper-ocean dynamics and thermodynamics. In CMIP6, the North Indian Ocean has shallow mixed layer depth biases in regions with thick observed barrier layers, driven by an easterly wind bias (Pang et al., 2024).

Ocean Heat Content

Ocean Heat Content (OHC) and its trends over time summarize the ocean's exchange of heat with the rest of the Earth System, and the associated buffering or amplification of atmospheric temperature changes. Globally over the 1970-2020 period, the CMIP6 OHC exhibits a warming bias (warmed faster than observations) near the surface (0 - 200 m) and below 2000 m depth, and a cooling bias between approximately 400 - 2000 m depth (Sohail et al., 2021). The opposing biases compensate and disappear when averaged over 0 - 700 m depth (Takano et al., 2023). This bias is attributable to inaccurate SST and biased mixing/surface fluxes into the warmest regions of the ocean, and the bias could be reduced through improved mixing

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

parameterizations (Sohail et al., 2021). In the Arctic, CMIP6 models exhibit a warm and fresh bias at depth through 1979-2014, with effectively no improvement in Arctic deep hydrography from CMIP5 to CMIP6. It was suggested that these Arctic biases could result from inaccurate mixing schemes, insufficient model resolution, and propagation of biases from adjacent ocean regions (Khosravi et al., 2022). The spatial structure of OHC biases in CMIP6 show a warm bias at 500-1000 m depths, particularly in the South and Equatorial Atlantic regions, and cold biases near the surface, particularly in and around the Arctic (Figure 9.7 of IPCC AR6 WG1 Chapter 9; Fox-Kemper et al., 2021). Regional CMIP6 upper ocean (0-700 m depth) OHC trends over 1971-2014 exhibit cooling biases across the equatorial North Atlantic and central West Pacific, and warming biases in the North Pacific. Across a deeper layer (0-2000 m) and shorter time period (2005-2014) CMIP6 OHC trends also exhibit regions of warming and cooling biases, with distinct spatial patterns relative to the 0-700 m depth 1971-2014 biases (Figure 9.6 (c) and (f) of IPCC AR6 WG1 Chapter 9; Fox-Kemper et al., 2021).

3.1.2 Sea Ice

Sea ice interacts dynamically and thermodynamically with the ocean and the atmosphere. At high latitudes, it plays a dominant role in ocean-atmosphere turbulent heat, momentum and gas exchange by forming a barrier between the two components. It also defines the albedo (reflectivity) of the ocean surface and thereby modulates the strength of radiative (short-wave) heating. A loss of sea ice decreases the albedo, increases the heat uptake of the ocean and therefore, leads to enhanced warming in polar regions, especially in the Arctic where sea ice exists at higher latitudes and through the summer. A correct representation in coupled climate models is essential for predicting future climate conditions. However, substantial deficiencies exist in the coupled model representation of sea ice area (SIA) in both hemispheres, including biases in the seasonal cycle, spatiotemporal distribution, and its response to climate change.

Observational estimates of mean Arctic SIA are generally captured by the mean of the CMIP6 models, though with a slight overestimation from January to May and a large spread among individual model climatologies throughout the year (Watts et al., 2021; Shen et al., 2021). There is no significant improvement in the seasonal cycle of Arctic SIA from CMIP5 to CMIP6. Single CMIP6 models show even larger biases than CMIP5 models, with overestimated SIA of up to $8 \cdot 10^6 km^2$ during boreal winter (Shen et al., 2021).

Model representation of the Antarctic SIA seasonal cycle is similarly spread around observations, with some extreme outliers producing $10 \cdot 10^6 km^2$ more or less than observed winter ice (Schroeter, 2025). The spread between model climatologies is wide enough that peak winter ice in some models is lower than the summer minimum in others (Roach et al., 2018). The majority of models tend to underestimate mean Antarctic sea ice, even during the austral winter, and fewer than half of the models fall within the observational SIA range throughout most of the year (Schroeter, 2025; Beadling et al., 2020). Moreover, almost the entire model cohort slows or delays the freezing of ice in autumn, which is linked to reduced winter ice cover, and a large proportion of models also delay the onset of spring sea ice melt (Schroeter, 2025). Strong asymmetry is a feature of the Antarctic sea ice seasonal cycle (Goosse et al., 2022), likely caused by insolation (Roach et al., 2022); however, it is poorly captured in approximately a third of the CMIP6 models, with some models producing equal-sized seasons instead, while others exaggerate the existing asymmetry with a four-month retreat phase, leading to rapid summer melt (Schroeter,

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

2025). Substantial regional biases also exist across the model cohort (Casagrande et al., 2023; Schroeter, 2025). Confidence in the current model representation of mean sea ice is further limited by the lack of realistic processes that govern sea ice cover in some models, in which sufficient mean ice can result from error compensation of competing inaccurate physical processes (Holmes et al., 2024; Nie et al., 2023).

Regarding the response of SIA to anthropogenic climate change, some distinct differences have been observed between the Arctic and Antarctic. The Arctic September sea ice extent has been declining significantly with a trend of -8.4 % per decade during the period of 1979 to 2005 (Meier et al., 2007). Matveeva and Semenov (2022) report an even more accelerated loss in September SIA in the period of 2000-2019 compared to 1979-1999. The Sea Ice Index from EUMETSAT OSI SAF and EUMETSAT SAF on Ocean and Sea Ice (2026), shows a mean decrease in September sea ice extent over the satellite period (1979-2025) of about 76,800 $km^2/year$, which represents a decrease of about 11.3 $\%/decade$ when referenced to the long-term (1981–2010) mean September sea ice extent of 6.88 million km^2. In the Antarctic, on the other hand, the annual SIA increased from 1979 to about 2015, only starting to show a negative trend after 2015 (Parkinson, 2019) (discussed in detail below). How well do the CMIP models capture these observed responses of sea ice in the Arctic and Antarctic to global warming? Notz and Community (2020) compare the sensitivity of Arctic September SIA to changes in global annual mean surface temperature (GMST) over the period of 1979 to 2014 for three different generations of CMIP models (CMIP3, CMIP5 and CMIP6). They find that CMIP models show a decrease in SIA with global warming, but underestimate the amplitude of September sea ice loss per °C warming. Although some improvement was found in CMIP5 and CMIP6 as compared to CMIP3 (see their Fig. 1d), a general underestimation of the sea ice sensitivity to changes in GMST remains for the multimodel mean of the latest CMIP model generation. However, while in CMIP3 there wasn't a single model falling into the 2σ plausible range, which Notz and Community (2020) defined including internal variability and observational uncertainty, 9 out of 40 CMIP5 and 11 out of 40 CMIP6 models fall into this range. Notz and Community (2020) conclude that it remains unclear whether this improvement is due to a better representation of sea ice processes in the latest generation of models or whether it is simply due to the improvement of the historical forcing (Fyfe et al., 2021). More recent observations indicate a slowdown in Arctic sea ice decline over the last few years (e.g., England et al., 2025), highlighting the importance of internal variability on observed sea ice trends. The representation of dSIA/dGMST in CMIP models would likely improve by considering the longer period of observations now.

A striking disparity exists between simulated and observed Antarctic sea ice trends for much of the satellite record. While observed Antarctic SIA expanded from 1979-2014 before undergoing a swift reversal and decline (Parkinson, 2019), simulated ice trends generally produce declining sea ice cover over the entire historical period (Hobbs et al., 2016). Several studies have reported varying mechanisms contributing to the observed expansion (Eayrs et al., 2019; Yu et al., 2023; Holmes et al., 2024)). However, in the years following the ice expansion of 1979-2014, an abrupt decline has occurred in the recent decade which is also regarded as a phase shift of the sea ice (Purich and Doddridge, 2023; Hobbs et al., 2024; Kusahara and Tatebe, 2025). When this trend component is included in comparisons with historical CMIP simulations, the long-term observed trend lies within the spread of modeled SIA trends, leading some to argue that models do in fact reasonably capture multidecadal Antarctic SIA trends overall (Holmes et al., 2024). In general, however, though there is little statistical relationship between observed SIA

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

and GMST, CMIP models tend to produce a negative correlation between the two, adding uncertainty to projections of future change Roach et al. (2020). Additional complexities of model-observation comparison include limitations in the robustness of existing satellite data (particularly in coastal areas and regions of low-concentration ice) (Meier et al., 2022), which constrain valuable tools such as sea ice budget analyses (Keen et al., 2021). Understanding of the sensitivity of annual Antarctic SIA loss per °C warming, therefore, remains limited. Moreover, the currently available models are extremely unlikely to simulate extreme events, such as the Antarctic sea ice loss in 2023 (Diamond et al., 2024).

3.1.3 Atlantic Meridional Overturning Circulation (AMOC)

The Atlantic Meridional Overturning Circulation (AMOC) is a key component of the global climate system, transporting heat and freshwater between the tropics and high latitudes. It strongly influences North Atlantic climate, SST variability, and large-scale feedbacks involving the atmosphere, sea ice, and carbon cycle (Fig. 3). Despite progress from CMIP5 to CMIP6, the AMOC remains one of the most biased large-scale ocean diagnostics in coupled models.

On average, CMIP6 models simulate a mean AMOC transport at 26°N of around 16–18 Sv, close to observations from RAPID array (17 Sv; Moat, 2025) but with a wide spread, ranging from ∼10 to 23 Sv (Weijer et al., 2020; Jackson et al., 2022). Menary et al. (2020) showed that CMIP6 models simulate a stronger and more realistic mean AMOC than CMIP5, mainly due to improved subpolar surface density and convection. However, they still produce a too-shallow North Atlantic Deep Water (NADW) cell and a weak, inconsistent link between AMOC strength and Labrador Sea density, revealing persistent biases in deep-water formation and variability.

Analyses across model generations reveal a persistent underestimation of low-frequency AMOC variability in CMIP models. Yan et al. (2018) showed that CMIP3 and CMIP5 models simulate too-weak multidecadal AMOC fluctuations compared with observations, and more recent assessments indicate that this deficiency persists in CMIP6. Zhao et al. (2024a) found a large diversity in the amplitude and timescales of internal AMOC variability among NEMO-based CMIP6 simulations, linked to differences in salinity-driven density variability and water-mass transformation in the Nordic Seas. Similarly, Reintges et al. (2024) showed that the mean-state temperature and salinity biases (See Sect. 3.1.1) in the subpolar North Atlantic strongly modulate the AMOC's response to the North Atlantic Oscillation, with warmer-saltier mean states producing a stronger and more delayed overturning response, highlighting the role of mean-state errors in shaping modelled AMOC variability (Fig. 3). In addition, Robson et al. (2023) demonstrated that CMIP6 models underestimate low-frequency AMOC variability and its coupling to the Atlantic Multidecadal Variability (AMV), resulting in an overly stable AMOC and reduced decadal predictability of the North Atlantic climate. Overall, coupled climate models simulate an overturning circulation that is overly stable, too shallow, and insufficiently sensitive to buoyancy forcing, a combination that may damp both internal variability and the response to external perturbations.

Beyond the variability bias, structural errors in the mean overturning circulation remain substantial. Heuzé (2021) found that CMIP6 models show a slightly improved representation of NADW properties compared to CMIP5, linked to more realistic surface density structures. However, most models still produce NADW that is too warm, too shallow, and too diffuse, indicating persistent deficiencies in deep convection and overflow representation. These biases lead to an overly diffuse deep limb and an

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

underestimation of the southward heat and freshwater transport (Weijer et al., 2020; Menary et al., 2020). Beyond mean-state biases, overturning pathway structure also matters: Baker et al. (2023) showed that inter-model differences in the connectivity between upper and lower overturning limbs strongly control the magnitude of projected AMOC weakening in CMIP6, indicating that circulation geometry is a first-order source of uncertainty in future projections. Heuzé (2021) also showed that CMIP6 models still struggle to represent Antarctic Bottom Water formation, with overly warm and fresh deep waters and insufficient bottom-water export, highlighting persistent global deficiencies in deep-water mass formation and circulation (Fig. 3). Some of these biases can be traced to unresolved overflow entrainment across the Greenland-Scotland Ridge (Bryden et al., 2024), insufficient eddy exchanges along continental margins (Swingedouw et al., 2022), and broad and diffusive boundary currents around Greenland that disperse runoff and meltwater unrealistically (Devilliers et al., 2021; Martin and Biastoch, 2023; Schiller-Weiss et al., 2024). Together, these findings suggest that structural differences among coupled models may represent a larger source of AMOC uncertainty than parametric choices alone, with important implications for the robustness of future projections.

The absence of an interactive Greenland Ice Sheet (GrIS) component prevents CMIP models from capturing realistic freshwater flux trends and their variability (Goelzer et al., 2020; Devilliers et al., 2021, 2024). Freshwater transport biases further complicate the simulation of their interaction of those fluxes with the AMOC. The boundary current around Greenland is often too broad and diffusive, dispersing runoff and sea-ice meltwater unrealistically across the subpolar gyre, weakening stratification and deep convection (Swingedouw et al., 2022; Schiller-Weiss et al., 2024). The transport of freshwater across 34°S is also systematically biased in CMIP models, with several even simulating the wrong sign of the net freshwater flux (van Westen and Dijkstra, 2024; Van Westen et al., 2024). This reverses the expected salinity feedback in the South Atlantic, disrupting the interhemispheric salt balance that helps sustain the AMOC. Such errors can weaken the salt-advection feedback and affect the equilibrium stability of the overturning circulation, potentially reducing the models' sensitivity to freshwater perturbations.

AMOC errors are closely linked to other coupled biases in the North Atlantic sector (Fig. 3). The persistent cold SST anomaly in the subpolar gyre, excessive sea-ice extent in the Labrador and Nordic Seas, and an overly fresh surface layer all act to suppress convection and limit deep-water formation (Fox-Kemper et al., 2021). Precipitation and surface freshwater flux biases further amplify these errors, while a misrepresented subpolar gyre circulation displaces the North Atlantic Current and distorts the distribution of heat and salt between subtropical and subpolar regions (Koul et al., 2020; Biastoch et al., 2021). The inadequate resolution of dense overflows across the Greenland-Scotland Ridge and weak eddy-driven exchanges along continental slopes (Hirschi et al., 2020; Martin and Biastoch, 2023) contribute to a sluggish and overly stable AMOC, hindering a realistic response to buoyancy and freshwater forcing.

The implications of these biases extend throughout the coupled system. These persistent deficiencies raise a fundamental question about the reliability of future projections. McCarthy and Caesar (2023) argue that models unable to reproduce the observed twentieth-century AMOC evolution should be interpreted cautiously when assessing future weakening, emphasizing the need for emergent constraints and process-based evaluation. An underestimated AMOC variability limits simulated AMV amplitude and predictability, weakening modeled teleconnections with tropical rainfall, African monsoons, and Northern Hemisphere temperature variability (Yeager and Robson, 2017; Robson et al., 2023). The cold SST bias south of Greenland influences atmospheric storm tracks, while excessive stratification and sea-ice cover inhibit surface heat exchange, maintaining

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

the bias (Weijer et al., 2025). An accurate representation of AMOC is also essential for transports and uptakes of anthropogenic CO_2 by the ocean and the subsequent interior ocean acidification (Brown et al., 2021; Tjiputra et al., 2023).

Reducing AMOC-related biases in future CMIP generations will require improved representation of key processes. Higher resolution (0.25°-0.5°) has already been shown to yield more realistic boundary currents, overflows, and gyre dynamics, resulting in improved NADW properties and stronger overturning (Martin and Biastoch, 2023; Hirschi et al., 2020). Large sensitivity to model formulation further complicates bias reduction. Roberts et al. (2020) demonstrated that relatively small changes in parameterizations, particularly mixing and overflow representations, can produce markedly different AMOC states, underscoring the role of structural uncertainty in shaping both the mean circulation and its variability. Enhanced parameterizations of convection and entrainment are also essential to better represent deep-water formation at CMIP-class resolutions. Controlled experiments that increase horizontal or vertical resolution, switch overflow parameterizations, or apply targeted freshwater perturbations can help identify which aspects of the AMOC bias are driven by model structure versus parameter settings. Besides, more process-oriented tests and targeted MIP experiments designed to separate structural biases from parametric choices will improve the understanding of the source of structural errors. Incorporating interactive ice-sheet and freshwater components will allow more realistic GrIS-ocean coupling and freshwater routing, while improving the treatment of surface fluxes, particularly precipitation, evaporation (See Sect. 3.3.2), and runoff (See Sect. 3.2.1), will reduce uncertainty in buoyancy forcing (Goelzer et al., 2020; Fox-Kemper et al., 2021).

Progress will also depend on improved benchmarking against observations. Expanded use of sustained observing systems such as RAPID (26°N; Moat, 2025), OSNAP (Lozier et al., 2019), and SAMBA, together with hydrographic programs such as OVIDE and Argo-based estimates, provides valuable constraints on AMOC strength, structure, and variability. Emerging observationally constrained frameworks further help reduce uncertainty in overturning estimates and model evaluation (Bonan et al., 2025; Portmann, 2025). Planned model configurations include improved surface freshwater budgets and coupling schemes with interactive cryosphere components. The inclusion of enhanced observational constraints and targeted process MIPs focusing on convection and overflows will help separate true model biases from observational uncertainties and enable a clearer evaluation of AMOC sensitivity and stability.

3.1.4 ENSO (lifecycle/seasonality/teleconnections)

El Niño Southern Oscillation (ENSO) is a coupled ocean-atmospheric phenomenon, characterised by tropical Pacific SSTs oscillating between anomalously warm (El Niño) and cool (La Niña) phases interspersed by years with SSTs close to the climatological average (neutral phase) (Santoso et al., 2023; Trenberth, 2020). ENSO is the leading mode of interannual climate variability and has significant impacts on regional temperature and rainfall patterns (Taschetto et al., 2020; Lieber et al., 2022) as well as extreme weather (Goddard and Gershunov, 2020; Lieber et al., 2024), leading to severe economic and societal impacts (Smith and Ubilava, 2017; Liu et al., 2023; McPhaden et al., 2006). It is also a critical Earth system component that can exacerbate key tipping elements such as Amazon rainforest dieback, west Antarctic ice sheet melting and coral reef bleaching (Wunderling et al., 2024). Addtionally, it has significant downstream impacts on the global carbon cycle, e.g. through

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

regulating the ocean carbon uptakes, which could feed back on the projected future climate change (Vaittinada Ayar et al., 2022) Simulating ENSO is a tricky endeavour, and the fact that this complex, nonlinear phenomenon emerges from the coupling of independently developed atmosphere and ocean components of climate models is a testament to the progress the modelling community has made since the first generation of GCMs (Guilyardi et al., 2020). With significant improvements made to atmosphere and ocean component models as well as coupling algorithms (Guilyardi et al., 2020; Bellenger et al., 2014; Planton et al., 2021a), all latest generation of coupled climate models now exhibit ENSO variability in the tropical Pacific. But, despite the efforts in improving ENSO representation and tropical Pacific dynamics, there are still long-standing biases in ENSO behaviour (Bellenger et al., 2014; Planton et al., 2021a, b; Guilyardi et al., 2020; Hou and Tang, 2022; Liao et al., 2021) which can be categorised broadly across four main categories:

– ***ENSO characteristic biases*** These are the key ENSO diagnostics that are used for evaluating ENSO behaviour and representation in models. An example is the *ENSO SST pattern*, which represents the SST anomaly in the central and eastern equatorial Pacific, which oscillates between its two active warm and cool phases. This oscillation typically has an amplitude of 1 °C which represents the *ENSO variability*. Furthermore, the ENSO SST pattern can be stronger for El Niño events compared to La Niña events, giving rise to *ENSO asymmetry* (Bayr et al., 2024). The peak of ENSO variability typically occurs during boreal winter, making ENSO seasonally phase-locked, and this characteristic is referred to as *ENSO seasonality* (Chen and Jin, 2022; Liao et al., 2021). Following its peak, an active ENSO phase can last for a few months to a year, which determines *ENSO duration*. The transition from one active ENSO phase to the opposite can be quick or can be interspersed by few ENSO neutral years, or in some cases even multiple occurrences of the same ENSO active phase in subsequent years (Freund et al., 2024), giving rise to a complex *ENSO lifecycle*. These characteristics are shaped by the background state and the balance between positive (Bjerknes) and negative (heat flux and cloud) feedbacks (Bayr et al., 2019).

– ***ENSO teleconnection biases*** ENSO SST anomalies in the tropical Pacific can have significant wide-ranging impact on rainfall and temperature patterns in far away regions, which are referred to as *teleconnections* (Taschetto et al., 2020). They arise as a result of large-scale reorganisation of atmospheric and ocean circulation processes due to the SST modulation in the tropical Pacific and cross-basin influences. The teleconnection patterns can also be asymmetric and nonlinear across active phases (Sengupta et al., 2025), with impacts often being more prominent for one phase compared to another.

– ***ENSO processes related biases*** ENSO is driven by key ocean-atmospheric feedback mechanisms – warm SST anomalies in the tropical Pacific can weaken trade winds, causing the thermocline to flatten, stratifying the upper ocean and suppressing cold water upwelling, which can then further enhance the warm SST anomalies. This positive feedback mechanism is called the Bjerknes feedback, which is often damped by negative heat flux feedback mechanisms – warm SST anomalies can increase low cloud cover which can decrease warming by reducing the incoming shortwave radiation and damping the warm SST anomalies (Planton et al., 2021a; Xiang et al., 2012; Priya and Dommenget, 2025). These positive and negative feedback mechanisms, or processes also operate in the opposite sense for negative SST anomalies.

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

They are complex and challenging to simulate due to their coupled nature and interactivity between the ocean and atmosphere, but also due to challenges in simulating finer sub-grid scale processes, such as mesoscale eddies in the ocean and cloud microphysics in the atmosphere, which can alter the magnitude of these feedbacks (Vijayeta and Dommenget, 2018).

– ***Mean state and seasonal cycle biases*** These are the biases in the background mean state and seasonality of the tropical Pacific temperature and rainfall patterns. Some examples of these are the long-standing cold tongue SST bias (Section 3.1.1), the "double ITCZ" and "dry equator" biases (Sections 3.3.2 and 3.3.3), and the mean state thermocline slope bias in the equatorial Pacific.

These categorical biases occur in all coupled climate models, but they are not independent of one another, and are often inter-related (Fig. 3). For example, biases in the background mean state can give rise to biases in ENSO-related processes and feedbacks, which can alter key ENSO characteristics. For instance, cold tongue bias in the equatorial Pacific and warm SST bias in the south-east Pacific can alter wind stress patterns which can give rise to Bjerknes feedback biases that can alter ENSO SST pattern, variability (Bayr et al., 2019, 2024) and eventually teleconnection strength (Planton et al., 2021a). Subtropical cloud biases in the south-east Pacific can alter heat flux damping feedbacks, which can reinforce the tropical mean state and ENSO pattern bias (Xu et al., 2025). Seasonal cycle biases in tropical convergence or zonal winds can alter the ENSO seasonal phase-locking and peak amplitude (Freund et al., 2020; Liao et al., 2021; Chen et al., 2022). Even GCMs that simulate ENSO characteristics well, often do so due to the wrong reasons, as a result of error compensation between positive and negative feedback mechanisms (Guilyardi et al., 2020; Planton et al., 2021a, b). This error compensation can give rise to correct ENSO SST pattern and amplitude but can cause large biases in second-order ENSO characteristics such as seasonality, lifecycle, and asymmetry. In addition to these key characteristics, models also struggle to capture 1) the diversity of ENSO events, i.e. ratio of events that peak in the Central Pacific (CP ENSO) to those that occur in the East Pacific (EP ENSO), often related to the SST biases (Santoso et al., 2015; Hou and Tang, 2022; Freund et al., 2020) and 2) the decadal scale variability in the tropical Pacific (discussed further in section 3.3.3). A systematic bias related to the ENSO lifecycle exists in several coupled climate models. It is associated with too weak spectral power at 2–7 years when compared to observations (Planton et al., 2021a; Bellenger et al., 2014). Typically, models are more biennial in their ENSO lifecycle, highlighting an incorrect representation of ENSO dynamics.

ENSO teleconnections are an essential metric to evaluate as they have regional impacts, and model projected changes in these impacts often inform policy decisions (Beniche et al., 2025; Lieber et al., 2022). While CMIP models typically perform well in simulating the linear nature of ENSO teleconnections, they often struggle to capture the asymmetric component of the teleconnections. More recent assessments of CMIP6 models suggest that the same models, which show a correct linear ENSO-rainfall relationship, struggle to capture the asymmetric component of this teleconnection (Sengupta et al., 2025). Regional drizzling biases and wet biases can significantly alter the local rainfall skewness, with climatologically dry regions in observations, often being "too wet" in coupled climate models. As a result, the rainfall response during each active phase of ENSO is then skewed in the wrong direction, causing a bias in the nonlinear component of the ENSO-rainfall relationship. Biases in regional rainfall impacts of ENSO are also linked to the model biases in ENSO diversity and event sequencing (Freund et al., 2024; Santoso et al., 2015; Ferrett et al., 2020; Chung et al., 2023). Central Pacific and Eastern Pacific ENSO events can have different impacts on regional rainfall patterns, and the sequence in which they occur can significantly alter regional drought development and also influence the probability of seasonal drought-breaking rainfall. Biases in general circulation such as the Walker circulation strength, Hadley circulation width and subtropical jets also strongly modulate regional teleconnection biases and are less well understood (Lee et al., 2021a; Bayr et al., 2019; Jiang et al., 2022).

ENSO process and feedback biases are the most complex and distinct across models and can emerge from limitations in atmospheric convection schemes (Bayr et al., 2019, 2024), which can influence zonal trade winds and tropical convection thresholds; cloud and boundary-layer parameterizations, which can shape surface heat flux feedbacks by modulating shortwave and longwave radiation; ocean vertical mixing schemes that can affect the thermocline slope; representation of mesoscale eddies and tropical instability waves (TIWs), which are often missing in coarse resolution models, and can reduce nonlinear damping of SST anomalies; and coupling frequency in flux exchange algorithms which can distort ENSO seasonality (Planton et al., 2021a).

The suite of CMIP6 models have shown improvements in ENSO pattern and teleconnections, compared to CMIP5, owing to a better representation of the mean state (Bellenger et al., 2014; Planton et al., 2021a; Brown et al., 2020; Hou and Tang, 2022). However, these biases are not completely alleviated. In addition, some key characteristics, like ENSO diversity, asymmetry, seasonality and even amplitude have shown little improvement, due to persistent process related biases across CMIP generations (Guilyardi et al., 2020; Planton et al., 2021b). Due to the complexity of ENSO feedback processes, and their error-compensating influence on ENSO diagnostics, it is difficult to make targeted model improvements, but continuing the improvements in systematic mean state biases are critical and already a key focus for CMIP7 (Dunne et al., 2025). In addition, developing high-resolution ocean and atmosphere models and improved subgrid parameterizations is critical to better represent finer-scale ocean and atmosphere processes, particularly instability waves, atmospheric convection, and low-level cloud feedbacks. For medium to low-resolution models, next-generation convection schemes, improved boundary layer and ocean mixing parameterizations are critical. Mean state correction approaches, such as flux adjustments, bias aware spin-ups and nudged initialisations, can also be employed before coupled runs to stabilise the background climate. Encouraging coordinated high-resolution experiments in community MIPs associated with CMIP7 and standardised ENSO diagnostic model benchmarking using existing tools such as the CLIVAR ENSO package (Planton et al., 2021a) should be primary objectives going forward.

Evaluation of ENSO diagnostics is complicated due to the lack of reliability in the observed SST records pre-1950, strong unforced variability of ENSO, and sampling uncertainties. Higher order ENSO characteristics such as skewness, diversity, nonlinear teleconnections can require multi-century samples for robust evaluation. Therefore, ENSO evaluation should be carried out using multiple different observation datasets with initial condition large ensembles to meaningfully constrain model ENSO variance and higher-order characteristics. Modelling groups should be encouraged to run multiple ensembles, particularly for low resolution runs, to fully incorporate sampling effects for benchmarking key ENSO diagnostics (Lee et al., 2021a; Maher et al., 2025).

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

3.2 Land and Land Ice

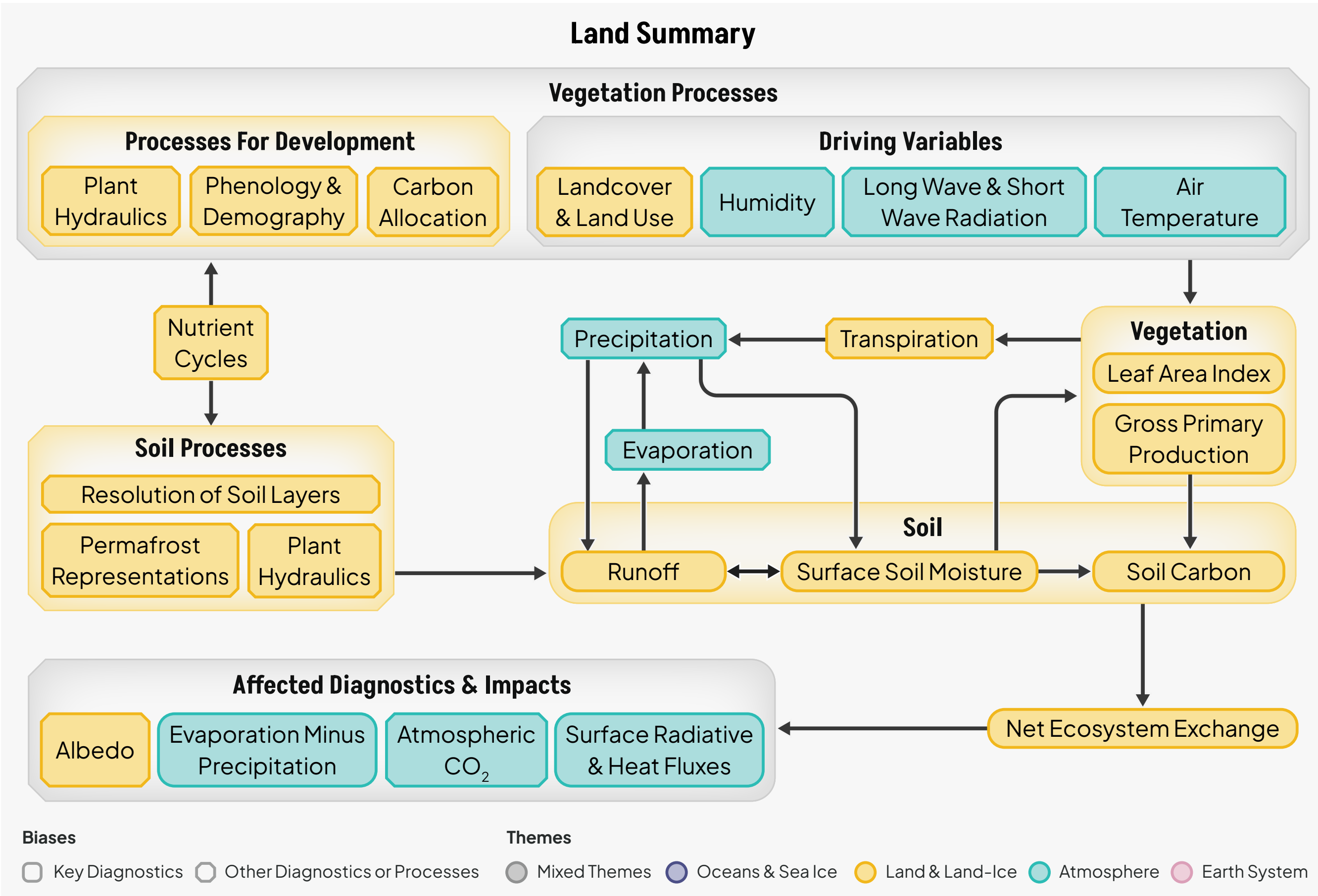

Figure 4. Key diagnostics and drivers of bias on the Land

3.2.1 Hydrology

Runoff and soil moisture are critical land surface variables central to understanding hydrology, land-atmosphere interactions, and moisture fluxes, and are therefore widely used to evaluate and constrain the hydrological realism of CMIP coupled climate models and offline land-surface models.

Runoff

Runoff biases remain a persistent weakness of coupled climate models and their embedded land surface schemes, with consequences that propagate from global water-cycle assessments to basin-scale impact studies (Hou et al., 2023; Guo et al., 2022). Evaluations of CMIP5 and CMIP6 simulations show that many models systematically overestimate mean annual runoff in

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

humid and monsoonal basins and in several large rivers, while also tending to overestimate runoff in arid and semi-arid basins such as the Murray–Darling and Nile, with substantial inter-model spread in magnitude (Guo et al., 2022; Miao et al., 2023). At the same time, global assessments reveal common deficiencies in the intra-annual cycle, most notably an overly large seasonal amplitude and an early peak in simulated runoff relative to observations, indicating errors in both the timing and volume of simulated streamflow (Hou et al., 2023).

Runoff biases in CMIP models arise from coupled errors in both atmospheric forcing and land-surface processes. At the forcing level, biases in precipitation (amount, intensity, persistence and seasonal evolution) are a primary driver of runoff errors, with global experiments showing that precipitation and temperature biases explain most of the variance in simulated runoff errors (Hou et al., 2023; Papadimitriou et al., 2017). Structural deficiencies in land-surface schemes, including overestimated evapotranspiration, simplified soil and vegetation parameterizations, and misrepresented snow accumulation and melt, further distort the partitioning of water between evapotranspiration, storage and runoff and can strongly modulate the sign and magnitude of projected runoff changes (Mueller and Seneviratne, 2014; Zhou et al., 2023). Recent studies further demonstrate that constraining CMIP projections with observations or emergent hydrological relationships tends to reduce ensemble spread while often amplifying projected regional runoff declines, underscoring the need to diagnose and reduce runoff biases to obtain reliable water-resource projections (Lehner et al., 2019; Kim et al., 2026). These findings motivate continued efforts to attribute and reduce runoff biases in CMIP models, including through improved land-surface representation, observational and emergent-constraint methods, and systematic benchmarking in forthcoming CMIP7 experiments (Dunne et al., 2025; Hoffman et al., 2025).

Surface Soil Moisture

Surface soil moisture directly impacts drought assessment, climate projections, and land-atmosphere coupling. Compared with CMIP5, the CMIP6 ensemble shows reduced biases in soil moisture simulations when compared to NLDAS-2 and in-situ datasets (Yuan et al., 2021), GLDAS (Sang et al., 2021), ERA5 (Dutta and Maity, 2022), and Station and ESA-CCI data (Wang et al., 2022a). Model evaluations across Eurasia and China also show higher pattern correlation and lower RMSE relative to observations (Sang et al., 2021; Wang et al., 2022a), indicating reduced biases at the regional scale. Additionally, the representation of regional soil moisture climatology has improved, especially in consistently wet areas and monsoon regions (Dutta and Maity, 2022). These improvements mainly result from enhanced land physics, including addition of new hydrology models, improved vertical soil layering with stronger coupling in transitional climates, snow and freeze-thaw processes, vegetation, and evapotranspiration (ET) parameterizations (Berg and Sheffield, 2019; Talib et al., 2023) (Fig. 4). CMIP6 captures long-term soil moisture trends well across all models, such as the overall drying trend, and shows high consistency in trend signs (Wang et al., 2022a). Over the contiguous United States, significant improvements in capturing wetting trends in the Midwest, Northeast, Southeast, and Northwest US are noted in CMIP6, an improvement over CMIP5 (Yuan et al., 2021). Moreover, the difference in CMIP5 and CMIP6 is spatially extensive, as both have consistency in representing large-scale patterns of wetting and drying. However, they disagree over transitional regions with robust drying and wetting responses, especially when the region switches from wetting to drying in CMIP6 (Cook et al., 2020). Projection analysis indicates that CMIP6 historical

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

simulations exhibit higher correlations with observations than CMIP5 simulations, suggesting improvements in the models' baseline performance (Jensen et al., 2020). For the near-present period, the bias in soil moisture trends relative to in-situ observations has been reduced by ∼40%; the inter-model spread in soil moisture trends has decreased as well (Yuan et al., 2021). Despite these advancements, challenges still exist. Many CMIP6 models still lack complex hydrological processes (e.g., lateral groundwater flow, lateral flow of reinfiltrating river water, and irrigation with river water), which contribute to persistent uncertainties (Hosseinzadehtalaei et al., 2023). A wide inter-model spread in soil moisture projections and coupling strength, and a systematic underestimation of short-term soil moisture-atmosphere feedback, add to the existing uncertainties (Talib et al., 2023). Refining the modules of hydrological processes (groundwater, lateral flow, irrigation) further, with more soil types, layers, and depths, for example, a five-layer soil hydrology scheme from the MPI-ESM1-2-HR from CMIP6 that divides the soil into distinct layers (top layer, root zone, and deep soil layer) (Jensen et al., 2020), may provide mitigation. Moreover, the propagation of bias from precipitation (see Sect. 3.3.2) to soil moisture remains a significant limitation, leading to higher uncertainty in future projections of soil moisture than in precipitation or temperature. The uncertainties lead not only to misrepresentation of soil moisture and associated feedbacks to the climate system, but also, prominently, to misrepresentation of drought risks. Development of global km-scale ESMs such as nextGEMS (Segura et al., 2025) could altogether reduce bias in precipitation simulations and model uncertainties stemming from parameterization schemes.

3.2.2 Ecosystems

Net Ecosytem Exchange

Net ecosystem exchange (NEE) is a key carbon cycle diagnostic representing the net carbon flux between the land and the atmosphere. It is an important component of the Transient climate response to cumulative emissions of carbon dioxide (TCRE) and is calculated as gross primary production (GPP) minus ecosystem respiration. Eddy covariance flux towers directly measure NEE and can be used to partition this exchange into GPP and respiration components (Desai et al., 2008; Li et al., 2025). However, ESMs typically simulate GPP and respiration processes separately from the bottom up, making NEE evaluation subject to compounded uncertainties from each carbon cycle process. Consequently, many modelling studies focus on GPP alone (Spafford and MacDougall, 2021). Net Biome Production (NBP) extends this framework by incorporating carbon losses from other land-use terms such as deforestation and fires. Gier et al. (2024) assessed NBP in CMIP5 and CMIP6 models, finding systematic biases (underestimation of uptake in the Northern Hemisphere and overestimation in the Southern Hemisphere) that offset each other to yield a seemingly accurate global mean. Improvement of spatial accuracy remains essential for understanding regional carbon dynamics; observations generally indicate positive NBP in the Northern Hemisphere, driven by limited deforestation and increasing afforestation, whereas the Amazon remains a major uncertainty due to ongoing deforestation and potential transition to a net carbon source. Current ESMs cannot resolve this question, as CMIP6 simulations exhibit a wider inter-model spread than observational uncertainties. Both NBP and NEE play central roles in linking the carbon and water cycles, e.g., via water-use efficiency. Accurate representation of these fluxes is thus critical for reliable carbon budget estimates. The inclusion of nitrogen cycle processes in several CMIP6 models is expected to improve scenario projections via improvement of GPP, given that the impact of nitrogen limitation on photosynthesis is anticipated to intensify in the future

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

(Fig. 4). Overall, because ecosystem fluxes represent the integration of multiple interacting processes, substantial uncertainties persist in both simulations and reference datasets, complicating bias analysis of NEE, NBP, and related ecosystem-level fluxes that underpin the global carbon cycle. This has discouraged substantial literature review of this bias.

Gross Primary Production

Gross Primary Production (GPP) is a key component of NEE. It represents the total carbon (C) taken up by an ecosystem through plant photosynthesis. Accurate estimation of GPP is critical for quantifying the terrestrial carbon sink and its role in removing CO_2 from the atmosphere. Additionally, GPP influences radiative budgets, surface heat fluxes and moisture fluxes (Fig. 4) through its control on stomatal conductance and evaporative fraction (Yuan et al., 2022). Reliable GPP representation is also essential for evaluating climate projections and the impacts of natural climate solutions, such as afforestation, on atmospheric CO_2 concentrations. In ESMs within CMIP5 and CMIP6, the latitudinal mean distribution generally aligns with observations but the magnitude tends to be overestimated (Gier et al., 2024). Compared to MODIS products this is by about 16% (Hu et al., 2022), although there is also substantial uncertainty and variability among the observation-derived products. Models especially overestimate GPP in areas of high productivity such as the tropics, but on average underestimate between 40°S-50°S (Hu et al., 2022). Between models, there is a considerable range in magnitude, and spatial biases also vary substantially (Hu et al., 2022; Gier et al., 2024). As a result, the multi-model mean (MMM) consistently performs better than individual simulations (Hu et al., 2022). CMIP6 models that included an interactive nitrogen cycle, which decreases GPP through nitrogen limitation, showed improved GPP magnitude relative to CMIP5, suggesting that this process should be incorporated in all models. Other developments to vegetation schemes did not result in significant improvements to GPP (Gier et al., 2024). Simulations still struggle to capture responses to climate variability, such as droughts and extreme events (Green, 2024), and the CO_2 response remains highly uncertain. The drivers of variability also differ widely between models, both in terms of contributing regions and dominant forcing factors (e.g., temperature, soil moisture, or precipitation), and limited empirical understanding hinders identification of missing processes or incomplete parameterizations (O'Sullivan et al., 2020; Dunkl et al., 2023). Several models aim for enhanced representation of phenology and demography to improve simulation of vegetation dynamics, whilst other studies suggest implementing hydraulic drought models that represent vegetation responses to water limitation. To better capture model sensitivity to climate forcing, several recent land surface model developments include temperature acclimation and CO_2 sensitivity, though additional empirical research is needed to assess their parameterization across ecosystems and species.

Leaf Area Index

Leaf Area Index (LAI) is the one-sided leaf area per unit ground area. This is essential for accurately estimating GPP, albedo, and carbon and water fluxes (Fig. 4). In CMIP5, LAI biases were identified as a major weakness of land surface modelling (Anav et al., 2013). Although CMIP6 shows improvements, with many individual models and the multi-model mean broadly reproducing observed spatial patterns of LAI, systematic biases remain (Song et al., 2021). All models show an overestimation of total LAI compared to observations, particularly in non-forested regions such as grasslands and tropical savannahs (Song

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

et al., 2021). In CMIP5, Anav et al. (2013) linked mid-latitude overestimations to a wet bias that produced excessive soil moisture and consequently excessive LAI. This issue is not fully resolved in CMIP6, although the inclusion of nutrient limitation has reduced the bias in some models (Gier et al., 2024). Additionally, around half of CMIP6 models underestimate LAI in tropical forests (Song et al., 2021). Models generally reproduce the expected spatial relationship of LAI with mean annual precipitation and mean annual temperature, with a few exceptions. In wet regions, models employing CLM land components tend to overestimate LAI, whereas other models with dynamic vegetation schemes more often underestimate it (Song et al., 2021). Biases are also evident in simulated temporal behaviour. Observed trends in LAI are underestimated in the Amazon and the Arctic but overestimated in the boreal forest, partly due to overly strong temperature sensitivity in the models. While many models capture the broad seasonal cycle across latitudes and regions, interannual variability is not captured by any model. In boreal regions, both CMIP5 and CMIP6 exhibit phenological errors, commonly delaying the end of the growing season, again likely linked to excessive temperature responsiveness (Park and Jeong, 2021).

LAI estimation depends on detailed vegetation processes that regulate carbon allocation, responses to water and temperature, and senescence. Some biases arise from errors in driving data (e.g., soil moisture, see Sect. 3.2.1), while others stem from incorrect plant functional type (PFT) distributions or inadequate representation of disturbances (e.g., fires). Nevertheless, a substantial portion of LAI bias is attributed to incomplete understanding and imperfect parameterization of carbon processes within vegetation schemes (Park and Jeong, 2021). Despite increases in spatial resolution and advances in vegetation and biogeochemistry modelling from CMIP5 to CMIP6, improvements remain uneven across models and regions. Enhanced representation of phenology remains a key requirement for reducing persistent LAI biases.

Soil Carbon

Critical to understanding the carbon cycle, soils represent the largest terrestrial store of carbon on the Earth's surface, containing at least twice the amount of carbon within global vegetation stocks. Due to their sensitivity to climate change, soils could become a significant carbon source in the future, counteracting the enhanced sink due to vegetation productivity (Varney et al., 2024). Soils play a vital role in ecosystems, respiring carbon back to the atmosphere via decomposition of organic matter and providing nutrients such as nitrogen for plants (Fig. 4). Spatially within ESMs, present day magnitudes of soil carbon stocks are underestimated within the northern high-latitudes; a bias which has remained between CMIP5 and CMIP6 (Varney et al., 2022; Todd-Brown et al., 2013). The systematic underestimation in the northern high-latitudes is often compensated by other regions, so the global totals are more consistent with observations; though a large uncertainty in observational datasets is noted. A likely reason for this bias is a lack of representation of northern high-latitude processes within CMIP ESMs, for example explicit permafrost physics. An improvement is expected in CMIP7 due to an increasing number of ESMs starting to better represent soil carbon in these regions, i.e. vertically-resolved soil (Burke et al., 2020). Some improvements were seen between CMIP5 and CMIP6, with two models including permafrost physics, however the bias remained in the ensemble mean. This bias results in a coupling within CMIP6 between soil carbon and Net Primary Productivity (NPP; the approximate input of carbon into the soil) which is not seen in observational data (Varney et al., 2022). This bias in turn results in consistent projections of soil carbon change in the future across the CMIP6 ensemble, due to a cancellation effect between the changing input and output

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

fluxes controlling soil carbon stocks, despite differences in the projected magnitudes of the individual fluxes between ESMs
605 (Varney et al., 2023). It has been shown that CMIP6 models underestimate the sensitivity of soil carbon to temperature within these northern high-latitude regions (Koven et al., 2017; Varney et al., 2020), which could potentially lead to an overestimation of the land carbon sink and inaccurate carbon budgets in the future.

3.3 Atmosphere

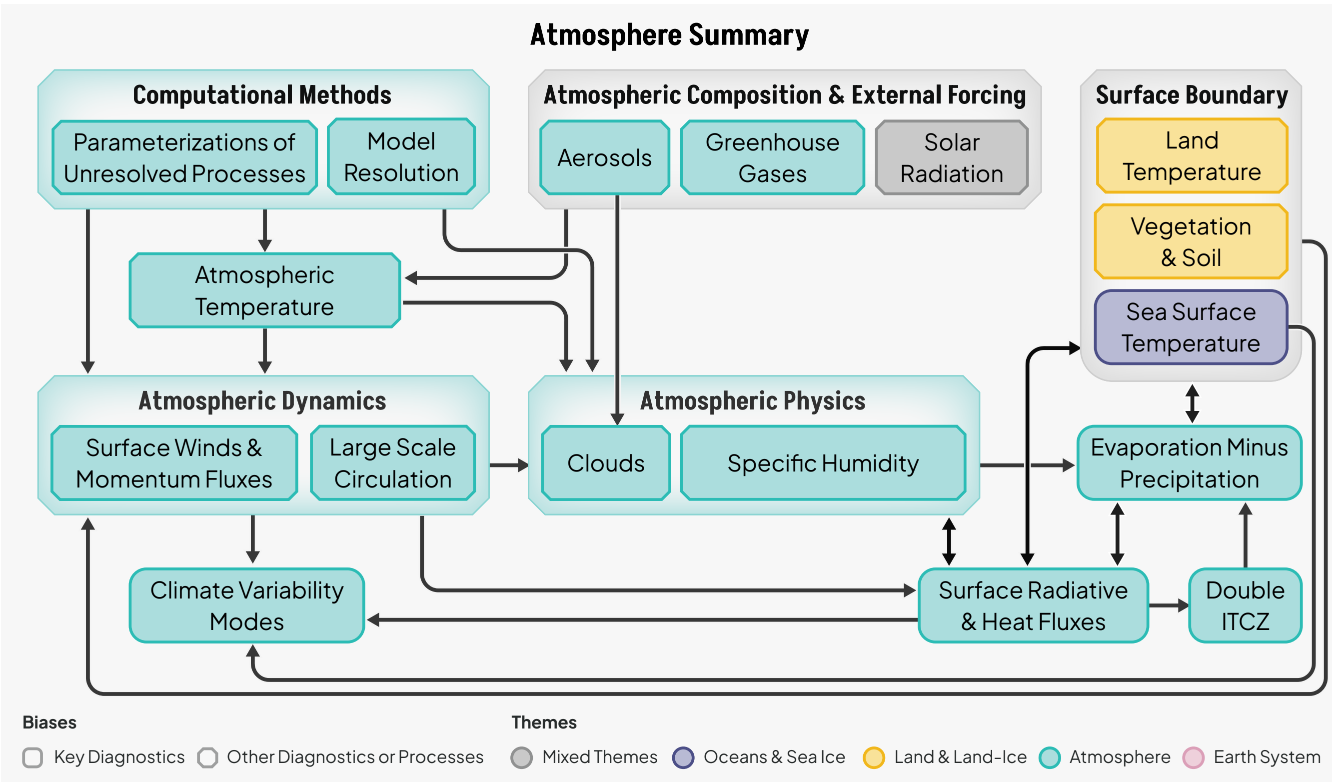

Figure 5. Key diagnostics and drivers of bias in the Atmosphere

3.3.1 Radiative Balance

610 **Radiative and Heat Fluxes at the Surface and TOA**

Surface radiation represents the radiative fluxes at the ground, which controls whether the surface warms or cools. Compared with direct observations, coupled climate models consistently overestimate the surface downward shortwave radiation and underestimate the surface downward longwave radiation, a systematic bias that persists across CMIP5 and CMIP6 (Wild et al., 2019; Wild, 2020). In particular, the surface downward shortwave radiation exhibits the largest positive biases over
615 extensive regions such as the Southern Ocean (SO), primarily due to persistent deficiencies in cloud representation (Bodas-

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Salcedo et al., 2008, 2014; Hyder et al., 2018). This overestimation of incoming solar radiation at the surface contributes to the well-documented warm SST biases in these regions (Fig. 5, Section 3.1.1), further amplifying coupled model errors in ocean-atmosphere interactions (Hyder et al., 2018). Compared with CMIP5, CMIP6 shows modest improvements in surface downward shortwave radiation over the Southern Ocean. These improvements are attributed to enhanced solar reflection due to improved representation of cloud fraction and phase (Cesana et al., 2022). At the global scale, CMIP6 exhibits a lower multi-model mean in surface absorption and a higher multi-model mean in atmospheric absorption, resulting in a reduced global mean of surface downward shortwave radiation (187.4 W m^{-2}) in CMIP6 compared to (189.6 W m^{-2}) in CMIP5. This reduction of $\sim$2 W m^{-2} brings CMIP6 into closer agreement with recent observational reference estimates (Wild et al., 2015, 2019; Wild, 2020). However, despite this improvement, substantial spread remains among the models in CMIP6, with global mean values of surface downward shortwave radiation differing by as much as 21 W m^{-2} (Wild, 2020). For longwave radiative fluxes at the surface on the global scale, the multi-model mean in CMIP6 indicates less effective surface longwave cooling, with a net surface longwave flux of -56.2 W m^{-2} compared to -58.6 W m^{-2} in CMIP5. This difference is primarily attributed to the higher surface downward longwave radiation in CMIP6, which exceeds that of CMIP5 by 3.7 W m^{-2}, along with a 1.2 W m^{-2} increase in surface upward longwave radiation in the CMIP6 multi-model mean. The higher global mean surface downward longwave radiation in CMIP6 brings the models into closer agreement with recent observational reference estimates (Wild, 2020). Overall, CMIP6 exhibits improved performance compared to CMIP5 in simulating both shortwave and longwave radiative fluxes at the surface. However, the substantial inter-model spread and persistent biases in surface radiative fluxes highlight the ongoing need for fundamental improvements in cloud and aerosol parameterizations. These challenges also call for dedicated efforts to systematically diagnose and characterize radiative processes within forcing frameworks (Dunne et al., 2025; Kramer et al., 2025).

Surface heat flux

Surface heat flux biases in CMIP6 refer to systematic errors in simulating surface energy exchanges, such as latent and sensible heat fluxes and radiative components, which regulate the global energy balance. These CMIP6 biases vary by region and seasons, with some improvements over CMIP5 but also persistent regional problems. For instance, CMIP6 models show a warm bias over the central United States in the summer due to issues in evapotranspiration (ET) partitioning and errors in surface downwelling longwave radiation (Dong et al., 2022; Sun and Liang, 2023). Similarly, latent heat flux biases higher than 10 W m^{-2} are observed over the Indian Ocean due to overestimated surface winds and vertical humidity gradients (Mohan and Ruchith, 2023). In general, errors in surface heat flux induce significant warm biases over land and tropical oceans, resulting from a combination of misrepresented physical processes and model parameterizations (Fig. 5). Land-surface processes, combined with underestimation of transpiration and overestimation of soil evaporation, result in negative ET biases and warm surface temperatures (Dong et al., 2022). Surface wind biases with the overestimation of wind speed also directly induce latent heat flux biases through excessive evaporation (Mohan and Ruchith, 2023). Likewise, uncertainties in cloud-aerosol interactions affect radiative fluxes and exacerbate biases in midlatitudes and polar regions (Zhao et al., 2024b). Aerosol forcing, particularly from SO_2 dry deposition also produces cold SST biases (Mulcahy et al., 2023). On the other hand, over the oceans, resolution,

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

among other factors, contributes to SST errors (Sect. 3.1.1). Cold SST biases in the northern Indian Ocean alter the strength of the South Asian monsoon, diminishing rainfall and moist convection (He et al., 2023). In contrast, dry and warm biases over land enhance soil drying and overestimate heatwave intensity (Al-Yaari et al., 2023). Similarly, feedback mechanisms are affected by biases in cloud cover and radiation, which influence SST variability and global precipitation patterns (Zheng et al., 2023; Exarchou et al., 2018). These shortfalls propagate into long-term projections, undermining confidence in estimates of temperature extremes, hydrologic variability, and monsoon circulation (Hu et al., 2023).

Cloud radiative effects

Cloud radiative effects (CRE) represent the impact of clouds on the global radiation budget, defined as the difference between all-sky and clear-sky radiative fluxes (Ramanathan et al., 1989; Li et al., 2026). CRE is governed by cloud fraction, morphology, and complex macro- and microphysical properties but remains a significant source of uncertainty across both CMIP5 and CMIP6 (Forster et al., 2021; Boucher et al., 2013; Zelinka et al., 2020; Li et al., 2024b; Stephens, 2005). Specifically, coupled climate models exhibit a persistent “too few, too bright” bias over tropical and subtropical subsidence regions and historically across the Southern Ocean, where models tend to underestimate cloud cover while overestimating cloud optical thickness (Nam et al., 2012; Klein et al., 2013). Furthermore, coupled climate models frequently exhibit an overestimation of total cloud cover north and south of the equator, a bias often associated with the “Double ITCZ” and deficiencies in convective parameterization (Oueslati and Bellon, 2015; Tian and Dong, 2020). While these biases remain a challenge in model assessment, CMIP6 shows notable progress compared to CMIP5 (Wild, 2020). This improvement is largely driven by a shift towards more sophisticated microphysics and higher supercooled liquid water fractions, particularly in the Southern Ocean, as well as the implementation of improved boundary layer and convection schemes that better represent cloud macro- and microphysical coupling (Zelinka et al., 2020; Hyder et al., 2018). As a result of these changes, at the global scale, the CMIP6 multi-model mean shortwave CRE is approximately -45.6 W m^{-2}, a slight intensification from the -44.8 W m^{-2} found in CMIP5, aligning more closely with recent observational estimates of ∼-46 to -48 W m^{-2} (Loeb et al., 2018; Wild, 2020). Conversely, the global mean longwave CRE in CMIP6 remains relatively stable at roughly 25.5 W m^{-2}, showing less variance between generations but still exhibiting significant regional biases in the tropical Pacific (Wild, 2020). Despite these ensemble-level improvements, the inter-model spread for net CRE in CMIP6 remains substantial, often exceeding 10 W m^{-2}, primarily due to differing treatments of cloud-aerosol interactions and convective parameterizations (Zelinka et al., 2020). At regional scales, notable discrepancies persist: while the underestimation of Southern Ocean shortwave cooling has improved between CMIP5 and CMIP6, persistent underestimations of shortwave CRE remain in subtropical stratocumulus regions, alongside an overestimation of both the shortwave and longwave CRE flanking the ITCZ (Lauer et al., 2023). These ongoing discrepancies are often rooted in structural limitations, such as coarse spatial resolution and observational uncertainties in polar and optically thin cloud detection ($\tau < 0.3$), highlighting a potential need for the CMIP7 framework to integrate kilometer-scale modeling, AI-driven parameterizations, and advanced active-sensing synergy from missions like EarthCARE to finally resolve the sub-grid-scale drivers of radiative bias (Lauer et al., 2023; Dunne et al., 2025).

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

3.3.2 Atmospheric Mean State

Annual Cycles & Seasonal Means

Many CMIP6 models exhibit substantial biases in representing the annual cycle of key climate variables, including precipitation and surface temperature, across diverse regions of the globe (Wang et al., 2020; Rivera and Arnould, 2020; Yan et al., 2020; Li et al., 2021; Hernandez and Chen, 2022). (Wang et al., 2020) identified significant and spatially coherent precipitation biases in CMIP6 simulations over the equatorial belt between 15°S and 15°N. Similarly, (Yan et al., 2020) reported that the CMIP6 multi-model ensemble (MME) substantially underestimates the amplitude of the observed precipitation annual cycle. Over South America, robust and regionally varying biases persist across many sub-regions, with coarse-resolution models in particular overestimating summer precipitation over southwestern South America, leading to pronounced seasonal wet biases (Rivera and Arnould, 2020; Almazroui et al., 2021).

Several CMIP6 models also simulate a delayed onset and extended withdrawal of major monsoon systems, accompanied by persistent wet biases over the North American monsoon region, although overall performance shows measurable improvement relative to CMIP5 (Hernandez and Chen, 2022). From CMIP3 to CMIP6, simulations of precipitation over India have improved through better representation of the annual cycle and a gradual reduction in the intensity and spatial extent of persistent dry biases over land and wet biases over the surrounding oceans (Choudhury et al., 2022). In the mid-latitudes, errors in the seasonal temperature cycle are frequently linked to deficiencies in the representation of snow-albedo feedbacks. Furthermore, only a limited subset of models realistically captures the annual evolution of eastern equatorial Pacific sea surface temperatures (SSTs) (Sect. 3.1.1) (Song et al., 2020). Collectively, these biases highlight the challenges that CMIP6 models still face in accurately capturing the annual and seasonal climate variations across different regions and climate regimes.

Evaporation minus Precipitation (E-P)

Evaporation minus precipitation (E−P) is a compact diagnostic for the net surface freshwater forcing of the climate system. Over the ocean it is essentially the atmospheric contribution to the surface freshwater flux that drives sea-surface salinity tendencies, upper-ocean stratification, and (via advection and mixing) broader water-mass properties. In steady state, E−P can be approximated by the vertically integrated moisture flux divergence ($\nabla \cdot \mathbf{F}_q$), so systematic E−P errors generally indicate biases in either the large-scale circulation that sets moisture transport (as the major agents for the transport; Hadley cell edges, Walker circulation, storm tracks) or in the physics that partitions moisture into precipitation versus boundary-layer export (O'Gorman and Schneider, 2008). This makes E−P useful for CMIP bias assessment because it "closes the loop" between (i) precipitation biases (e.g., misplaced convergence zones), (ii) evaporation/latent-heat-flux biases (surface winds, humidity gradients, air-sea coupling), and (iii) circulation biases (moisture flux divergence and its seasonality).

Many CMIP6 models reproduce the broad zonal-mean structure (subtropical net evaporation and tropical/high-latitude net precipitation), yet retain large regional E−P biases and strong inter-model spread, especially in the tropical oceans, subtropical dry zones, and the Southern Ocean. A persistent culprit is the double-ITCZ family of errors (Sect. 3.3.3): when precipitation is split into two bands or displaced meridionally, the tropical moisture budget and hence E−P gradients are distorted, with

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

knock-on effects for salinity patterns and coupled feedbacks (Tian and Dong, 2020). Model spread in projected $E-P$ has been shown to reflect a mix of thermodynamic amplification of existing patterns and dynamical/circulation-driven changes, with the relative contribution varying strongly by region and across models (Elbaum et al., 2022). This matters because a model can "look acceptable" in mean precipitation while still having compensating errors between E and P that yield the wrong freshwater forcing and the wrong coupling to SST, clouds, and circulation (Fig. 5).

E−P links atmospheric errors to ocean model drifts and large-scale ocean biases through the freshwater budget (Sect. 3.1.1). Several analyses of CMIP6 models highlight that imperfect conservation and model drift in mass/salt can complicate the interpretation of long coupled runs; separating physically meaningful E−P signals from spurious drift is therefore part of good practice when comparing models and observations or reanalyses (Irving et al., 2021). Regionally, biases in net freshwater forcing have been implicated in persistent North Atlantic water-mass and salinity errors, and in biases affecting high-latitude freshwater storage and export (Wang et al., 2022b; van Westen and Dijkstra, 2024). Regarding the bias assessment, a useful framing is therefore to treat E−P as a process connector, rather than "just another surface field". Therefore, it can be decomposed into contributions from circulation/moisture transport, surface turbulent flux biases, and precipitation physics, and evaluated jointly with sea-surface salinity, near-surface winds/humidity, and key rainfall error modes (e.g., ITCZ structure) to identify which fixes are most likely to reduce coupled-model bias.

Double Intertropical Convergence Zone (ITCZ)

The long-standing double-ITCZ bias remains a prominent deficiency in coupled climate model generations (Wang et al., 2020; Moreno-Chamarro et al., 2022), with a large inter-model spread persisting across CMIP3, CMIP5, and CMIP6 ensembles (Tian and Dong, 2020). Although the bias shows a modest reduction from CMIP3 and CMIP5 to CMIP6, and the annual equatorial Pacific cold tongue bias persists, the spread across CMIP6 models is narrower than in CMIP5. Notably, a sub-ensemble of the top-performing CMIP6 models has largely eliminated the double-ITCZ feature, an improvement not achieved in previous generations (Si et al., 2021). The origin of this bias is often attributed to systematic SST errors (Sect. 3.1.1) , particularly warm biases in the southeastern tropical Pacific and cold biases in the equatorial eastern Pacific, which promote excessive convection south of the equator and artificially shift the ITCZ. These errors are further reinforced by cloud and radiation biases (Sect. 3.3.1) and weak ocean-atmosphere coupling, whereas the models with minimal double-ITCZ signatures tend to exhibit very small warm SST biases in the southeast subtropical Pacific (Si et al., 2021). The presence of the double-ITCZ bias has wide-ranging consequences, contributing to large precipitation and monsoon biases globally and being closely linked to a shallow thermocline bias in the equatorial Pacific (Samuels et al., 2021). Higher model resolution reduces the double-ITCZ bias, but persistent tropical precipitation errors highlight the need for improved convection schemes or convection-permitting models (Moreno-Chamarro et al., 2022). Proposed pathways for reducing this bias include improved convective parameterizations, higher model resolution, and better simulation of SST fields; however, recent work suggests that increasing resolution alone does not substantially reduce the Pacific double-ITCZ bias (Ma et al., 2023). Because the ITCZ plays a central role in driving tropical-extratropical interactions, the double-ITCZ bias also propagates into errors in the seasonal cycle, monsoon rainfall,

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

temperature patterns, and large-scale circulation, linking this feature directly to broader annual-cycle and regional climate biases (Fig. 5).

3.3.3 Modes of Variability

Tropical modes of variability

ENSO is the primary tropical mode of climate variability, which has been extensively covered in section 3.1.4. Besides ENSO, there are other key tropical modes of variability across multiple time scales – Pacific Decadal Variability (PDV, Mantua et al., 1997; Capotondi et al., 2023; Power et al., 2021), Madden-Julian Oscillation (MJO, Wang et al., 2025) and Indian Ocean Dipole (IOD, McKenna et al., 2020).

PDV is a general term for the decadal-scale climate variability in the Pacific that encompasses both the Pacific Decadal Oscillation (PDO, Mantua et al., 1997) and Inter-decadal Pacific Oscillation (IPO, Power et al., 1999), which represent low-frequency shifts of the SST pattern in the Pacific Ocean basin. PDV is highly underestimated in CMIP models, with significant biases in both the SST pattern and the zonal winds (Kociuba and Power, 2015; Power et al., 2021) in the tropical and North Pacific that are key factors driving the SST variability on decadal timescales (Fasullo et al., 2020; Kajtar et al., 2018). The adequate simulation of PDV is vital for historical attribution, seasonal and decadal prediction, and multidecadal projections of global mean surface temperature and the pattern of future changes in SSTs. Internal fluctuations can fully obscure or amplify the underlying climate-change signal in many fields for decades, leading to periods of acceleration or hiatus in global warming (Maher et al., 2014; Douville and Cheng, 2024; Minière et al., 2023). Current literature suggests that the diminished decadal scale variability partly arises due to ENSO simulation in coupled models being too quasi-biennial and not persistent enough (given that long-term ENSO rectification plays a major role in PDV). Recent work also suggests that PDV biases in models are linked to Atlantic Ocean SST biases through inter-basin interactions (McGregor et al., 2018, 2014) involving displacements of key atmospheric pressure centres affecting the Walker cell and trade wind trends in the tropical Pacific. Suppressed variability in PDV has also been shown to be linked with suppressed South Pacific variability in models, due to alteration of coupled processes between the equatorial Pacific and South Pacific (Chung et al., 2019). Better representation of ENSO seasonality and mean state climatologies of SST are necessary to improve PDV representation in models, in addition to more process understanding of cloud and wind feedbacks in the tropics and extra-tropics, inter-basin interactions (Richter et al., 2025) and more reliable observations of low-level clouds over South-East Pacific and the Southern Ocean (Fig. 5).

The Quasi-Biennial Oscillation (QBO) is the dominant mode of variability in the tropical lower stratosphere and is active throughout the entire calendar year (Ebdon and Veryard, 1961; Reed et al., 1961; Lindzen and Holton, 1968; Baldwin et al., 2001). It consists of a nearly regular cycle of alternating eastward and westward winds that completes a full phase roughly every 28 months, with the wind regimes descending gradually from the stratosphere toward the upper troposphere at about 1 km per month. The QBO influences deep convection, as well as the thermal structure, humidity, and chemical composition of the equatorial stratosphere and upper troposphere (Collimore et al., 2003). It also modulates the Asian monsoon (Meehl, 1997) and modifies the propagation of planetary waves towards the high latitudes in what is known as the *Holton-Tan* effect, thereby influencing the strength of the stratospheric polar vortices, especially in the Northern Hemisphere (Holton and Tan, 1980). In

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

general, the westerly QBO phase tends to strengthen the polar vortex and mid-latitude jet stream and is associated with fewer sudden stratospheric warming events, while the easterly phase has opposite effects (Anstey and Shepherd, 2014).

The representation of the QBO in CMIP has improved considerably over time. While none of the CMIP3-class models was able to reproduce the oscillation, 5 out of 35 CMIP5 models and 15 out of 30 CMIP6 models were documented to provide a reasonable representation of its main characteristics in the stratosphere (Richter and Tokinaga, 2020). The mean QBO *period* is well reproduced in CMIP5 and 6, but the mean range is overestimated, particularly in CMIP6. The QBO *amplitude* is largely underestimated in the lower stratosphere in many models, meaning that they tend to displace the phenomenon upwards relative to observations. The *vertical extent* of the QBO is generally larger in the models than in observations. Vertical resolution is key for the successful reproduction of the interactions between atmospheric waves and the mean flow that drive the shear layers' downward propagation (Geller et al., 2016; Holt et al., 2022). Butchart et al. (2023) demonstrated that the use of the same ozone forcing dataset, including the observed QBO signal, is sufficient to synchronize the simulated QBOs from different CMIP6 models with each other and with respect to observations. This result is remarkable in light of the generally low signal-to-noise ratios found for the effects of single forcing agents on atmospheric dynamics (Shepherd, 2014).

Similar to PDV, MJO variability is suppressed in CMIP models, with large negative biases in the eastward propagation of the MJO over the Maritime Continent (Ahn et al., 2020). This largely stems from the incomplete simulation of processes at fine temporal and spatial scales, such as vertical and horizontal moisture advection, as well as dry surface temperature biases over the Indo-Pacific warm pool region (Le et al., 2021). The MJO plays an important role in modulating extreme rainfall events in some of the most densely populated parts of the world (Wang et al., 2025). Therefore, a better representation of the MJO is critical to understand how the MJO will change under anthropogenic warming and what consequences these changes will have on intra-seasonal extremes. The MJO is often better represented in models with higher resolution (km-scale models) (Segura et al., 2025), making resolution a core focus for future CMIP iterations.

The IOD is another critical tropical mode of variability that influences temperature and rainfall patterns in countries located in-and-around the Indian Ocean basin, and therefore influences some of the most important agricultural heartlands of the world (Gurazada et al., 2024; Reddy et al., 2021; Rustiana et al., 2019). Biases in IOD variability and mean state within CMIP models are heavily linked to biases in ENSO variability through inter-basin interactions (McKenna et al., 2020). Another major source of biases in IOD variability stems from mean state SST biases in the tropical west and east Indian Ocean linked to seasonal surface wind biases (McKenna et al., 2024). A better representation of Indian Ocean seasonality and its interaction with the Pacific Ocean basin are critical for better representing IOD variability.

Extratropical modes of variability

The primary extratropical modes of variability are the Northern and Southern Annular modes (NAM, SAM) in the Northern and Southern Hemisphere respectively. The North Atlantic Oscillation (NAO), Pacific/North American Pattern (PNA) and Pacific Decadal Oscillation (PDO) also occur in the Northern Hemisphere. These modes of variability exert significant influence on weather conditions locally, as well as remotely via teleconnections. CMIP5 models have been shown to generally reproduce the spatial patterns of these modes of variability, but tend to underestimate the amount of variance explained by these

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

modes (Lee et al., 2019). The spatial patterns of NAM, SAM and NAO are better reproduced, with the PNA typically less well reproduced (Lee et al., 2019; Coburn and Pryor, 2021). Biases in both spatial patterns and explained variance have generally improved between successive CMIP3, CMIP5 and CMIP6 simulations (Fasullo et al., 2020; Lee et al., 2021b). A key bias of extratropical modes of variability is related to seasonality, with the persistence of modes into transitional or post-dominant seasons (Lee et al., 2019). There is some evidence that the more accurate spatial patterns of some modes of variability are related to higher numbers of vertical levels (Coburn and Pryor, 2021).

Synoptic weather types

Weather types (aka weather regimes or circulation types) describe the atmospheric circulation on finer scales than the large-scale modes of atmospheric low frequency variability mentioned above (NAM, SAM, NAO, PNA and others). They typically operate on spatial scales covering a few thousand kilometers and on time scales ranging from hours to days (Jones et al., 1993; James, 2007; Huth et al., 2008). Describing the direction and strength of the atmospheric flow as well as the sign of the vorticity (cyclonic or anticyclonic), weather types are synoptic-scale processes that indirectly describe the physical properties of the advected air masses and also resolve extreme weather events. Thus, their climatological statistics, such as frequencies or transition probabilities, are closely linked to local-scale climate conditions and directly relevant for impact and adaptation studies. Indeed, it has been argued that statistical bias correction schemes should only be applied to model data if the model reasonably reproduces the observed weather type statistics (Maraun et al., 2017).

Due to their relatively small scale and large dependence on dynamical rather than thermodynamic processes, error attribution for weather types is hindered by atmospheric noise and small signal-to-noise ratios when trying to isolate the effects of convection, clouds, boundary layer and orographic drag on the overlying circulation (Shepherd, 2014). Nevertheless, large inter-model error differences have been reported in all CMIP phases (Hulme et al., 1993; Osborn et al., 1999; Perez et al., 2014; Otero et al., 2018; Cannon, 2020; Borato et al., 2024), with CMIP6 model versions performing better on average than those participating in CMIP5, a large imprint of the model families on the error characteristics, and smaller model errors favoured by the use of a finer horizontal and vertical resolution in the atmosphere model component (Brands, 2022a, b; Brands et al., 2023).

Atmospheric blocking and cut-off low episodes are two specific weather types that occur along the ridges and troughs of the geopotential height field drawn by the meandering polar jet stream. Both event types are favoured by High Amplitude Rossby Waves (see Section 3.5) and can cause hydrological extreme events of opposite sign. Persistent anticyclonic conditions block the zonal flow of moist maritime air masses and are associated with subsidence, adiabatic heating, and dryness, which help maintain the blocking and ultimately lead to severe drought events (Woollings et al., 2018). The blocking frequency in the Northern Hemisphere is systematically underestimated in historical coupled climate model simulations, but this bias has improved considerably from CMIP3 to 6 (Davini and D'Andrea, 2020). This has been attributed to an increase in model resolution (Jung et al., 2012; Davini and D'Andrea, 2016; Schiemann et al., 2017), a reduction in the North Atlantic SST bias (Scaife et al., 2011; O'Reilly et al., 2016), and improved tropical convection (Jung et al., 2010; Gollan et al., 2019).

Cut-off lows are extratropical cyclones originating from equatorward moving troughs of cold air in the mid-troposphere that detach from the polar jet stream (Nieto et al., 2005). These systems move slowly after detaching from the jet and can determine

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

the weather of the affected region for more than a week, eventually causing persistent moisture advection, heavy rainfall, flooding and landslide events, as occurred in August 2002 in central-to-eastern Europe and in October 2024 in the Valencia region (Calvo-Sancho et al., 2026). The few available model evaluation studies for this phenomenon indicate an underestimation of their intensity on both hemispheres that has considerably improved from CMIP5 to 6, primarily due to an increase in model resolution. The annual cycle of the occurrence frequencies and life-cycle characteristics are generally well reproduced by both model cohorts (Pinheiro et al., 2022; Mishra et al., 2025).

3.4 Earth System

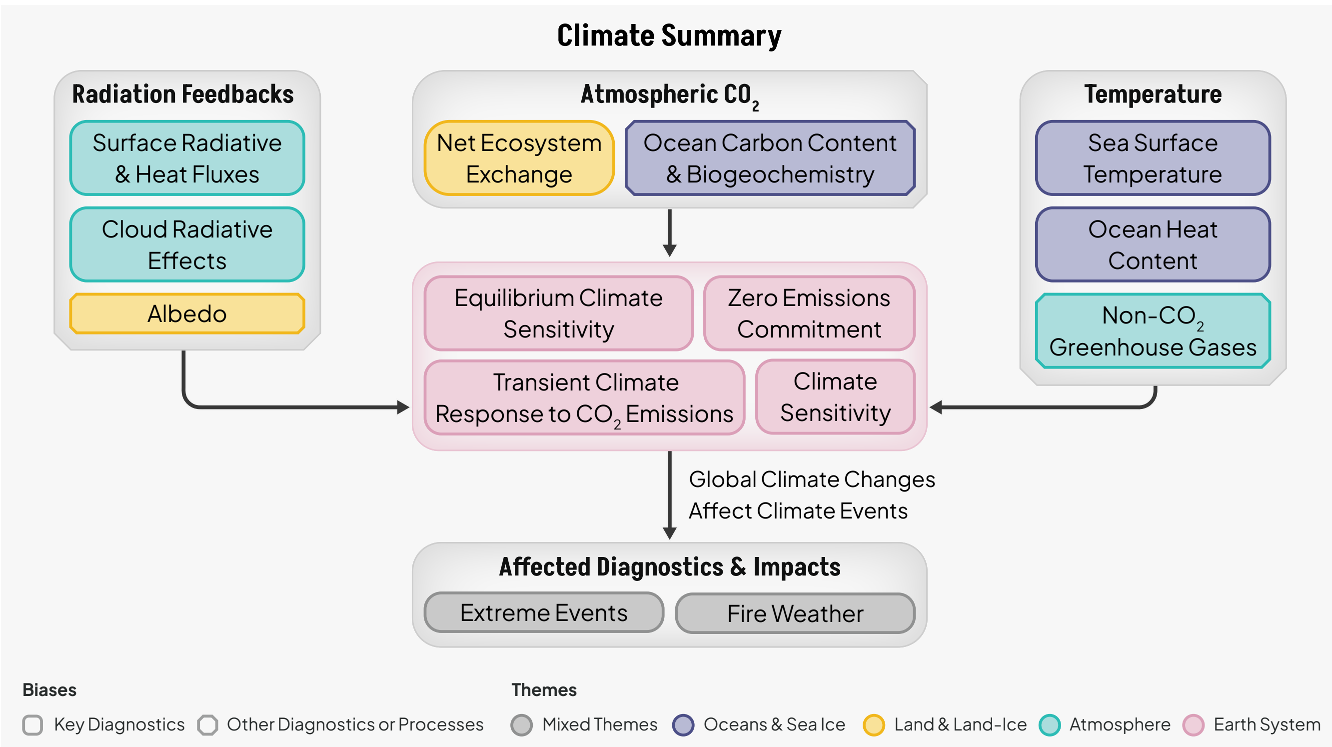

Figure 6. Key diagnostics and drivers of bias in the Earth System, and their contribution to Impacts and Adaptations

Exactly how much the Earth's surface will warm with rising atmospheric CO_2 concentrations is one of the most important questions in climate science, as it is an important constraint on greenhouse gas emission targets to avoid exceeding dangerous global temperature thresholds. A typical metric is to consider how much the Earth's surface warms in response to a change in the energy imbalance at the top-of the atmosphere. In principle, this simple metric is meant to capture the sum of all of the complex and competing feedback processes in the Earth system. The key feedback processes are related to clouds, surface albedo, water vapour, lapse rate and the Planck feedback, and the strength of these processes are expected to change over time. The following subsections will cover the 4 main ways to look at climate sensitivity (IPCC, 2021): Equilibrium Climate

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Sensitivity (ECS), Transient Climate Response (TCR), Transient Climate Response to Cumulative Emissions (TCRE) and Zero Emissions Commitment (ZEC). These different measures consider different timescales and different emission pathways. These diagnostics are considered as an uncertainty rather than bias because climate sensitivity is an abstract measure of the climate system so there is no direct observations to compare against, although observations can provide some constraints.

3.4.1 Climate Sensitivity

Equilibrium Climate sensitivity (ECS)

Equilibrium Climate sensitivity (ECS) is a measure of how much the Earth's surface will warm in response to a doubling of atmospheric CO_2 above pre-industrial levels once the climate system have reached equilibrium. While the central estimates of ECS have remained similar across different generations of models, the uncertainty has historically been difficult to constrain and even increased in CMIP6. The modelled range of ECS in CMIP3 was 2.1-4.4 °C, in CMIP5 2.1-4.7°C, and in CMIP6 widened to 1.8-5.6°C (Bock et al., 2020; Meehl et al., 2020). In CMIP6, a group of 8 (out of 50) models have particularly high values of ECS above 5. The existence of these very high ECS models, generated a lot of interest and research as they suggest that the climate may warm faster than previously thought.

Throughout CMIP3 to CMIP6, the major source of ECS uncertainty is linked to the long-standing uncertainty in cloud feedbacks (Bony et al., 2006; Vial et al., 2013), i.e. exactly how clouds respond to a warming climate which in turn affect the incoming solar radiation and impact surface warming (Fig. 6). For an overview of cloud biases see Sect. 3.3.1. Additionally in CMIP6, the primary factors causing higher ECS are stronger cloud-aerosol interactions and stronger cloud feedbacks (Meehl et al., 2020; Zelinka et al., 2020; Dong et al., 2020). Cloud changes over the Southern Ocean are thought to be particularly important (Zelinka et al., 2020). A puzzling aspect of this increased uncertainty is that some of these high ECS models show many improvements in related cloud biases, for example, improved mixed-phase clouds over the Southern Ocean (Bodas-Salcedo et al., 2019). These high ECS models tend to have overly strong aerosol forcing (Wyser et al., 2020; Andrews et al., 2019).

The realism of high ECS models in CMIP6 remain a topic of debate. Other sources of evidence, such as from paleoclimate comparisons (Zhu et al., 2020) and observational constraints (Sherwood et al., 2020), put the ECS at around 3°C as a central estimate, making the high ECS models unlikely. For this reason, high ECS models were excluded in the IPCC AR6 report for the assessed warming projections, the assessed range closer to the observational constraint than the CMIP6 range (IPCC, 2021).

Transient Climate Sensitivity (TCR)

Transient Climate Sensitivity (TCR) is a related metric but on a shorter timescale. TCR is defined as the amount the Earth's surface warms when CO_2 concentration is increased by 1% year-on-year until it is double its original amount (70 years). TCR tends to be lower than ECS because the ocean has not yetfully warm and continues to absorb heat. and many feedbacks are dependent on transient patterns of SST and are expected to change with time, additionally many feedbacks are dependent on transient patterns of SST and are expected to change with time (Armour et al., 2013; Andrews et al., 2019)

Although models with high ECS also tend to have high TCR, TCR is more constrained across the different generations of models compared to ECS. The TCR range for CMIP3 and CMIP5 remained 1.1-2.5 °C and only slightly increased to 1.3-3

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

°C in CMIP6 (Meehl et al., 2020). Additionally, observational estimates of TCR are in much better agreement with coupled climate models compared to ECS (Knutti et al., 2017) because TCR is determined by the ratio of observed warming to forcing. However, there is also evidence that for simulations extending to 2100 and beyond, ECS may be a better predictor, especially in high emission scenarios because it can better capture the effects of changing feedback strengths (Huusko et al., 2021).
The main difference between TCR and ECS therefore is that TCR is more relevant for near-term warming, whereas ECS is able to better capture the effects of delayed warming, of which large uncertainties remain. Although high climate sensitivity models are deemed as unlikely in AR6, the associated impact could be very high, therefore the risk of the actual world having a high ECS is still an important factor to consider (Sutton, 2019). Furthermore, regional impacts such as heavy rainfall events, droughts and fire, may be unrelated to climate sensitivity (Swaminathan et al., 2024). Therefore, neglecting high ECS models in regional impact and risk assessments may lead to an underestimation of climate risks. It has also been noted that, although there is still much work in understanding climate sensitivity, it's uncertainty should not be an obstacle for action on reducing CO_2 emissions (Knutti et al., 2017).

3.4.2 Emissions-related Biases

Projecting climate under emission-driven scenarios requires full ESMs capable of simulating the carbon cycle and its feedbacks, marking a key distinction from the concentration-driven experiments that dominate earlier CMIP assessments. As CMIP7 places greater emphasis on simulations where CO_2 is prescribed as an emission flux rather than an atmospheric concentration (Dunne et al., 2025), the ability of ESMs to accurately represent the coupling between carbon and climate becomes critical. Two diagnostics are particularly central to evaluating this capability: TCRE, which characterises the near-linear relationship between cumulative carbon emissions and global temperature change, and ZEC, which quantifies the temperature response following a complete cessation of emissions. Biases in either quantity have direct implications for carbon budgets and the credibility of net-zero pathways simulated by CMIP7 ESMs.

Transient climate response to cumulative emissions of carbon dioxide (TCRE)

The transient climate response to cumulative emissions of CO_2 (TCRE) quantifies the amount of warming per unit of cumulative emissions, representing a ratio between changes in global temperature to changes in atmospheric CO_2 concentrations (MacDougall, 2016). This response is governed by multiple, complex processes, however despite this, the ratio is found to be approximately constant over time and across all emission scenarios (Canadell et al., 2021). The proportionality between temperature and CO_2 emissions has allowed for a simplified metric representing the Earth system, used to calculate carbon budgets which are required for Paris Agreement policymaking (Cox et al., 2024). The uncertainty in underlying mechanisms is large and the need to constrain components of this response is required for an improved confidence in TCRE (Jones and Friedlingstein, 2020; MacDougall et al., 2020).
TCRE has been assessed in CMIP6 ESMs, where Arora et al. (2020) find the ensemble mean to align with the CMIP5 estimate, as well as additional estimates using lower complexity models and observational constraints (Canadell et al., 2021). However, uncertainties and biases within CMIP6 ESMs will be implicit within the TCRE estimates. TCRE is dependent on two main

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

components; the Transient Climate Response (TCR) and the airborne fraction of anthropogenic CO_2 emissions over time; both of which are subject to projection uncertainty. TCR is subject to biases within the physical climate system and how this responds to climate change. The fraction of CO_2 remaining in the atmosphere is due to the strength of the carbon cycle sink, both on the land (See Sect. 3.2.2) and ocean (Fig. 6). Biases in CMIP6 will therefore have subsequent consequences on the confidence of net simplified metrics used for policymaking.

Zero Emissions Commitment (ZEC)

ZEC quantifies the global mean surface temperature change we can expect following a complete cessation of anthropogenic CO_2 emissions (Palazzo Corner et al., 2023). The value of ZEC is typically calculated as a 20-year average after 25, 50, or 90 years post emissions cessation, but the estimation method depends on the experimental protocol. Based on estimates from CMIP6 models participating in ZECMIP, the multimodel average value of ZEC is 0 °C, but with substantial model spread in both the sign and magnitude (Jones et al., 2019; MacDougall et al., 2020). A more recent study by Borowiak et al. (2025b) suggests that the multimodel average is close to -0.19 °C with an uncertainty range of -0.44 °C to 0.04 °C, when filtering for variability on multiple timescales. This large inter-model uncertainty range has major implications for the remaining global carbon budget to stay below the Paris Agreement temperature target of limiting warming to well below 2 °C before 2100. Negative values of ZEC allow for more ambitious targets, while a positive value implies a reduction in the remaining budget, thereby influencing national and regional mitigation strategies (Palazzo Corner et al., 2023; Borowiak et al., 2025a; King et al., 2025). While a ZEC of 0 implies no net warming post net-zero CO_2 emissions, many other aspects of the Earth system such as sea ice extent, sea level rise (Palazzo Corner et al., 2023; Zickfeld et al., 2023), severe heatwaves (Perkins-Kirkpatrick et al., 2025; Cassidy et al., 2025, 2023; Borowiak et al., 2025a) and regional rainfall patterns (Dittus et al., 2024) will continue to evolve adversely. Constraining the value of ZEC is therefore critical from an impacts and adaptation perspective as well (Fig. 6).

The value of ZEC, in a simple approximation, depends on the energy balance in the Earth system post net-zero – if the Earth's surface and lower atmosphere absorb (emit) more energy than they emit (absorb) then ZEC would be positive (negative). This balance is heavily dependent on *atmospheric processes*, such changes in low-cloud decks; *cryospheric processes*, such as ice-albedo feedback, *physical & biogeochemical ocean processes* such as SST pattern effect, ocean circulation and ventilation changes; ocean stratification; and carbon and heat transfer; as well as *land processes* such as permafrost changes, future changes in wildfires, soil carbon feedbacks, nutrient controls on primary productivity, land use changes and its effects on carbon uptake efficiency. All of these processes are either not present in present-day coupled models or are highly uncertain in ESMs. Many of these uncertainties stem from the systematic biases discussed in the previous sections.

Certain model biases, such as ocean and land processes that influence carbon and heat uptake, need to be alleviated to better constrain ZEC. In CMIP7 fast track, several experiments have been proposed such as *esm-flat10-zec*, *esm-flat10-cdr* and ScenarioMIP7 ESM scenarios (Sanderson et al., 2024) that will be run in emissions-driven mode with a fully interactive carbon cycle. An inter-comparison of the results from these experiments is essential to understand the sources of uncertainties in ZEC. These uncertainties vary significantly on shorter decadal to longer millennial time scales (Palazzo Corner et al., 2023). There-

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

fore, extended millennial-length ESM net-zero simulations at different warming levels are critical (King et al., 2025, 2021) to better constrain ZEC uncertainties and gain an improved understanding of the underlying physical processes.

3.5 Impacts & Adaptation

High Amplitude Rossby Waves

High-amplitude Rossby waves (HARW) are unusually large and quasi-stationary oscillations in the mid-latitude jet stream that persist for multiple weeks, often manifesting as Rossby wave patterns of zonal wavenumber between 5 and 8 (Kornhuber et al., 2020; Luo et al., 2022). Such waves, which are highly amplified, can become circum-global, promoting stagnant weather regimes that lead to concurrent heatwaves, droughts, or heavy rainfall in widely separated regions (Petoukhov et al., 2013; Kornhuber et al., 2020). Observational analyses have shown that quasi-resonant amplification of planetary waves (wave numbers 6-8) underlies many recent summer extremes (e.g. the 2003 European heatwave, 2010 Russian heatwave/Pakistan floods, 2021 Pacific Northwest heatwave) (Petoukhov et al., 2013) and more recent work has examined how the occurrence of quasi-resonant amplification events may change under anthropogenic warming (Guimarães et al., 2024). In coupled climate models, however, these high-wave number wave phenomena remain challenging to present, particularly in terms of persistence and associated surface extremes (Luo et al., 2022). Both CMIP5 and CMIP6 simulations struggle to reproduce the observed frequency and amplitude of HARWs, often underestimating the occurrence of blocking and persistent ridges in summer (Davini and D'Andrea, 2020). CMIP6 models show some improvement (benefiting from higher resolution and better dynamics) compared to earlier generations, but substantial biases remain in planetary-wave-related circulation features (Davini and D'Andrea, 2020). For example, CMIP6 still underestimates summertime blocking frequency over Eurasia and fails to capture the recent increase in high-latitude blocking (e.g. over Greenland) (Delhasse et al., 2021; Maddison et al., 2024; Simpson et al., 2025). This systematic shortfall indicates a model bias (i.e., a consistent error in representing the mean climate behaviour of Rossby waves), rather than mere internal variability, and it undermines confidence in extreme event projections (Simpson et al., 2025). Several factors contribute to the poor simulation of high-amplitude Rossby waves. Limited horizontal resolution smooths out sharp jet-stream gradients and weakens the development of smaller-scale (high-wavenumber) waves, while biases in the mean jet state (e.g., jets that are too strong or displaced) alter Rossby wave propagation and breaking (Schiemann et al., 2017; Davini and D'Andrea, 2020). In particular, models may misrepresent the dynamical processes that maintain blocking (especially the interaction between wave breaking and transient eddy forcing) leading to overly progressive, fast-moving flow anomalies that do not amplify or stall as in reality (Woollings et al., 2018; Nakamura and Huang, 2018). Besides, transient eddies can feed back onto the large-scale anomaly by organising momentum and vorticity fluxes that reinforce a displaced jet and/or blocked ridge, hence, the strength and generality of this eddy feedback remains an active topic of research (Nakamura and Huang, 2018; Kautz et al., 2022). As a result, CMIP models tend to underestimate the amplitude of surface anomalies associated with HARWs: for instance, during observed high-amplitude wave episodes (i.e., an amplified 5/7 wave pattern), models typically produce temperature and precipitation extremes only about two-thirds as strong as in reanalysis (Luo et al., 2022). Such biases in wave amplitude, frequency, and phase speed reduce the reliability of climate projections for mid-latitude extremes and compound events. Reassuringly, planned developments for CMIP7 target many of these shortcomings. Higher-resolution at-

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

mospheres and improved physical parameterizations are expected to better resolve jet-stream variability and wave-mean-flow interactions, thereby enhancing the simulation of HARWs (Schiemann et al., 2017; De Luca et al., 2024). As the earlier studies show that increasing model resolution significantly reduces biases in blocking and large-scale circulation patterns, so the next generation of models should more robustly capture high-amplitude Rossby wave dynamics and their associated extreme weather, improving confidence in future projections.

Climate Drivers for Fire

Rising global temperatures and changing precipitation intensity and frequency are driving more frequent and severe fire weather. Wildfires have a positive climate feedback mechanism: not only do they devastate communities and habitats after ignition and rapid spread, but they also release vast amounts of carbon, and other GHGs, exacerbating climate change and dryness (Di Virgilio et al., 2019). Primary drivers of wildfires are high temperature, low precipitation, specific humidity, vapor pressure deficits, and strong winds (Abatzoglou and Kolden, 2013; Harris et al., 2020; Gaboriau et al., 2020), land use, vegetation type, and human activity linked to ignitions (Balch et al., 2017; Fernández-Guisuraga et al., 2021; Gallo et al., 2023). In particular, wildfire occurrence and the extent of burned areas (BA) are significantly associated with higher temperatures (Koutsias et al., 2013; Cardil et al., 2015), drought (Littell et al., 2016) and low precipitation (Flannigan and Harrington, 1988; Holden et al., 2018).

Examining specific climatic drivers of wildfire reveals notable differences in bias characteristics between CMIP5 and CMIP6. Bock et al. (2020) quantified progress across CMIP Phases 3, 5, and 6 and found that temperature, water vapor, and zonal wind speed display smaller biases in CMIP6 than in previous CMIPs. The global mean climatology for precipitation was found to be significantly above normal, but the global RMSE decreased significantly from CMIP3 to CMIP6. Conversely, precipitation biases associated with the double ITCZ (see Sect. 3.3.3) have persisted across all CMIP phases. Extreme precipitation, however, is better represented in CMIP6 than in CMIP5 (Abdelmoaty et al., 2021). Further research by Jiang et al. (2021) and Tian et al. (2024) showed that longstanding biases exhibited by clouds, temperature, specific humidity, and relative humidity across all atmospheric pressure levels since CMIP3 have been considerably reduced in CMIP6, with an important exception being relative humidity in the boundary layer. Additionally, CMIP5 models failed to capture tropical circulation systems, such as the ITCZ and the Walker circulation, with precision, leading to biased wind fields. These biases have been reduced in CMIP6, leading to noticeable improvements in the corresponding wind fields (see section 3.3.3).

Fire danger indices such as the Canadian Fire Weather Index (FWI) measure the compound meteorological fire danger derived from temperature, wind and humidity variables, and also consider the availability of fuel. Performance for the FWI was reported to be satisfactory for a number of models used in CMIP5, whose global temporal-mean FWI patterns match the observed one with pattern correlation coefficients between 0.7 and 0.9 for individual models, low centred spatial RMSEs, and a slight overestimation of the spatial standard deviation, as revealed by an analysis based on Taylor diagrams (Bedia et al., 2015). Gallo et al. (2023) found the CMIP6 model versions to perform better for these statistics, although a strict like-for-like comparison is not possible due to differences in the underlying reference datasets and study periods.

Apart from studies that evaluated specific climate drivers of fire weather, Kloster and Lasslop (2017) evaluated fire simulations

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

from CMIP5 and found that CMIP5 models underestimated BA by more than 50%, failed to reproduce its spatial pattern, and showed weak seasonal variability. In contrast, CMIP6 models showed significant improvements, with better representation of spatial patterns and estimates of BA and seasonality across regions where CMIP5 failed (Li et al., 2024a). The CMIP6 ensemble also shows a larger model spread in BA and in how combustion completeness is treated for fire emissions. These advancements are primarily due to shifting the fire model from GlobFIRM to the Li fire parameterization scheme (Li et al., 2013). The last decade noted a significant decline in BA due to agricultural expansion (Andela et al., 2017; Feng et al., 2025). However, this decline in BA and the corresponding fire carbon emissions are not reproduced by CMIP6 models, primarily because they do not consider the human influence on the availability of fire fuel (Li et al., 2024a). The CMIP6 ensemble still underestimates the global BA, but shows improvement over CMIP5, partly due to models not simulating fires lasting multiple days, which can lead to larger fires and increased BA (Li et al., 2024a). Another suggested factor is the larger Equilibrium Climate Sensitivity (ECS) of CMIP6 models compared to CMIP5 (See section 3.4.1), meaning the Earth warms more per unit of CO_2 increase (Zelinka et al., 2020). Rising temperatures allow the atmosphere to hold more water vapor, thereby increasing evaporation and evapotranspiration rates. Consequently, poor representation of vapor pressure deficits can lead to underestimation of BA in model simulations.

Existing biases in fire drivers propagate directly into fire simulation, increasing the uncertainty in BA and fire emissions. Many models use simplified fire schemes that poorly represent human ignitions, suppression, and land management. Additionally, coarse spatial resolution limits the simulation of fuel distributions and topographic effects. Coupling among wildfire emissions, aerosols, large-scale wind patterns, and climate feedback warrants further exploration, especially regarding plume rise and smoke-cloud interactions. Lastly, but importantly, uncertainty in global ground observations further constrains robust validation of simulated burned area and emissions.

4 Discussion

4.1 Common drivers among biases

Systematic biases arise from interacting uncertainties and parameterizations governing the exchange of energy, water, and carbon across the Earth system. These biases are not confined to individual components: they typically originate in radiative forcing and cloud processes, propagate through the atmospheric circulation, are modified by land and ocean dynamics, and are amplified by biosphere feedbacks and large-scale climate variability. As CMIP models share common structural assumptions, such as similar cloud microphysics schemes, coarse spatial resolution, and simplified land and ocean biogeochemistry, many biases are persistent across generations of models.

Ocean processes are a significant source of persistent CMIP6 bias because they control long-term heat and carbon uptake (Fig. 6). Coarse horizontal resolution limits representation of mesoscale eddies, boundary currents, and coastal upwelling, which regulate heat and salt. Parameterized mixing affects ocean stratification and deep-ocean heat uptake, contributing to widespread SST and salinity biases (Fig. 3), including the equatorial Pacific cold-tongue bias and Southern Ocean warm bias seen across CMIP6 models. Bias in wind stress, freshwater fluxes, and river discharge influences salinity structure and density gradients,

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

thereby affecting deep-water formation and the strength of overturning circulation. These biases propagate into large-scale climate modes, such as ENSO and the AMOC (Fig. 3), influencing global precipitation and temperature patterns.

Systematic biases in sea ice area (SIA) in both the Arctic and Antarctic arise from the spatiotemporal distribution of ice cover, and the response of sea ice to anthropogenic warming. These discrepancies are associated with deficiencies in the representation of mass balance components, ice-albedo feedbacks, and the underlying physical processes that control sea ice cover. The omission or oversimplification of key processes in some models, such as ice-sheet-ocean interactions and freshwater fluxes, further reduces confidence in simulated sea ice dynamics and projections, and feed into biases in the ocean (Fig. 3).

In the case of land-surface processes, parameterization of plant functional types, phenology, nutrient limitation, and disturbance regimes leads to oversimplification of vegetation and biogeochemical processes in CMIP6 which can bias estimates of LAI, GPP, and evapotranspiration (Fig. 4). Many models still rely on simplified soil hydrology schemes and missing subgrid heterogeneity in soil texture and topography (Fig. 4). This leads to systematic errors in soil moisture persistence, evapotranspiration, and land-atmosphere coupling strength, which in turn affect heat-wave intensity and drought projections. Despite a consensus on importance of permafrost in the global climate system under climate change, many CMIP models still lack specialized land surface schemes that realistically capture permafrost dynamics (Matthes et al., 2025). Although a number of studies have used CMIP models to investigate present-day permafrost characteristics, such as permafrost extent, mean annual near-surface temperature, and active layer depth (Chadburn et al., 2017; Burke et al., 2020), with respect to simulating permafrost extent and active layer depth, the performance of CMIP6 remains comparable to that of CMIP5 (Matthes et al., 2025). As a result, biases arising due to lack of representation of permafronst dynamics, snow processes, soil freeze-thaw cycles, and permafrost thaw also affect the representation of water availability, contributing to biases in high-latitude temperature, carbon-cycle, and emission estimates. Limited representation of vertically resolved soil carbon pools, microbial decomposition, and carbon-nutrient coupling affects soil carbon storage and its temperature sensitivity, leading to a potential overestimation of the land carbon sink, affecting projections of future carbon budgets (Fig. 6). Anthropogenic land-use change, irrigation, fire management, and groundwater exert substantial control over surface energy balance, hydrological processes, and terrestrial carbon exchange, and current representations contribute significantly to uncertainties in estimates of NEE and NBP.

Radiative forcing and cloud processes remain primary sources of CMIP6 biases (Fig. 5). Aerosol forcing and atmospheric composition are major drivers: uncertainties in aerosol-cloud interactions, aerosol optical properties, and historical emissions introduce spread in effective radiative forcing estimates across CMIP6 models. Inadequate representation of cloud fraction, phase partitioning, and optical depth affects both incoming solar radiation and the trapping of outgoing longwave radiation, contributing to biases in land and ocean, such as warm biases in the Southern Ocean (Fig. 3). On the other hand, biases in surface exchange from land and ocean, especially via albedo, influence high-latitude feedbacks and land-ocean temperature contrasts (Fig. 5). These radiative biases distort temperature gradients, introducing biases in pressure systems, atmospheric circulations, and precipitation patterns. Many biases in cloud and circulation representation arise from model resolution and parameterization. Surface winds, mixing and turbulence schemes affect low-cloud formation and moisture transport, reinforcing cloud-radiation biases. Convection schemes often misrepresent entrainment rates, trigger thresholds, and convective organization, contributing to the double ITCZ problem, unrealistic monsoon variability, and shifted storm tracks. As large-

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

scale circulation depends on temperature gradients and latent heat release, these atmospheric biases distort major modes of variability such as ENSO, IOD, and NAO (Fig. 5), affecting regional climate projections.

Crucially, these biases do not remain localized but interact across scales and domains. Structural limitations arising from finite model resolution and imperfect parameterizations manifest heterogeneously across different components of the Earth system. As a result, fine-scale uncertainties introduced at the parameterization level can cascade into large-scale systematic errors, ultimately affecting projections of future warming, carbon budgets, and the timing and magnitude of climate responses (Fig. 6). For ECS, uncertainties are dominated by radiative feedbacks, particularly through parameterization of cloud microphysics, cloud fraction, and convection (Fig. 5). TCR is additionally sensitive to ocean heat uptake efficiency, which is strongly influenced by biases in ocean circulation and mixing. Inadequate representation of mesoscale eddies and vertical mixing leads to errors in the rate at which heat is transported into the deep ocean, thereby affecting the transient warming trajectory (Fig. 3). Consequently, models with similar ECS may exhibit different TCR values due to differing representations of ocean processes and associated heat uptake. TCRE integrates both physical climate sensitivity and carbon cycle feedbacks, making it particularly sensitive to biases in land and ocean carbon uptake. Uncertainties in processes such as permafrost thaw, soil carbon decomposition, vegetation dynamics (Fig. 4), and ocean biogeochemistry influence the fraction of emitted CO_2 that remains in the atmosphere (Fig. 6). In high-latitude regions, uncertainties associated with permafrost thaw can also affect estimates of non-CO_2 GHG emissions, especially methane. ZEC depends on the balance between declining ocean heat uptake and continued carbon cycle feedbacks after emissions cease. Biases in ocean thermal inertia, circulation changes, and carbon sequestration processes (e.g., deep ocean uptake and land carbon persistence) affect whether warming stabilizes, continues, or reverses. In particular, errors in representing slow feedbacks such as permafrost carbon release or deep ocean equilibration introduce significant uncertainty in long-term committed warming.

4.2 Key literature gaps and observational uncertainty

This review provides a model performance overview across various CMIP phases for the most relevant diagnostics of the global climate system, based on those currently considered by the REF. However, snow cover, internal variability and climate variables at Global Warming Levels are not part of this review paper. The broad scope of the 'internal variability' and the 'climate variables at global warming levels' diagnostics restricted a simple literature review, although we highlight that these will be valuable inclusions to the REF. Additionally, historical changes in climate variables are not reviewed as an individual bias but are discussed with each diagnostic, where relevant.

Many of the ocean and atmosphere biases summarized in this manuscript are well-reviewed for CMIP6 models. While many aspects of the land surface are also covered by the literature, greater uncertainty or missing observations meant these diagnostics had fewer literature sources. In particular, for many land diagnostics the source of the bias remains unidentified, in part due to simplified process representation in models that makes attribution more challenging. Many diagnostics remain weakly constrained by observations, particularly those requiring fine-scale or subsurface observations. In the atmosphere, cloud microphysics, aerosol-cloud interactions, and convective processes lack direct observational constraints. In the ocean, there are limitations in observing subsurface structure and mesoscale to submesoscale dynamics. Many cryospheric processes, including

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

sea ice thickness, ice sheet dynamics, and permafrost thaw, are similarly under-constrained, especially over long timescales. Where observations are available, there is not necessarily high confidence surrounding the bias as some diagnostics have uncertainty in their reference datasets. Different observational datasets (e.g., satellites, reanalyses, in situ networks) can produce substantially different estimates of the same variable. Many 'observed' global-scale products in fact undergo some processing, therefore introducing uncertainty depending on, for example, the retrieval algorithm used. In other cases, diagnostics are derived from multiple observations (e.g., carbon budgets) compounding uncertainty, or cannot be observed at all. Diagnostics that are not directly observable must be inferred from observations of related variables, often combined with modelling. Site measurements can provide a direct observation when global measurements are not available, however this also has uncertainty associated with how representative the point measurement is of the model gridbox average, with greater uncertainty if high heterogeneity is expected within the gridbox such as complex orography or urban/rural edges.

Differences in temporal characteristics also affect model evaluation, particularly for processes operating on sub-daily timescales that are often highly relevant for impacts. Variables such as rainfall and wind are typically observed at high temporal resolution but are aggregated to daily scales for comparison with CMIP outputs, potentially obscuring important process-level differences that influence the aggregated signal. Furthermore, there exists a degree of near-term irreducibility in observational uncertainty, especially in the detection of long-term trends. Multi-decadal variability limits the ability to robustly identify externally forced signals and evaluate them in models without sufficiently long observational records. As a result, even with expanding observational systems, confidence in long-term changes, particularly for slow components such as ocean heat content, remains limited.

Several variables regularly feature in this review as contributing to biases in key diagnostics, but are not included in the list of selected REF diagnostics (Hoffman et al., 2025). Both ocean carbon cycle and biogeochemistry play a key role in regulating ocean carbon uptake and therefore carbon-climate feedbacks that influence the Earth system diagnostics (Sect. 3.4). Biases in atmospheric winds and broader large-scale circulation have a fundamental influence on ocean dynamics, air-sea feedbacks, regional climate patterns, and contribute to systematic circulation related errors in CMIP models. Non-CO_2 greenhouse gases (e.g., methane) affect radiative forcing and, also relevant to radiative forcing budgets, surface albedo would be a particularly useful diagnostic for high-latitude and cryospheric regions. Within the impact and adaptation theme (Sect. 3.5), extremes in temperature and precipitation are not directly considered, even though extreme events are the main drivers of climate risk. Incorporating these diagnostics into future evaluations or into the REF would contribute to a more comprehensive understanding of CMIP performance.

5 Conclusions and perspectives on CMIP7 and beyond

CMIP7 is driven primarily by four central research questions concerning *patterns of SST change*, *evolving weather and climate extremes*, *the water-carbon-climate nexus*, and the *risk of tipping points in the Earth system* (Dunne et al., 2025). A major expected development is a move toward higher spatial resolution in both atmospheric and ocean model components. Finer spatial scales are likely to improve representation of mesoscale eddies, tropical instability waves, and cloud microphysics. These developments are expected to alleviate some long-standing structural biases (Moreno-Chamarro et al., 2022). However, im-

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

provements in large-scale features such as the double ITCZ, AMOC strength, ENSO asymmetry, etc. will depend critically on how increased resolution interacts with parameterized physical processes, rather than the resolution alone (Kroll et al., 2025). CMIP7 is also expected to feature refined parameterization schemes. While some models are expected to have high-resolution components that may approach convection-permitting or eddy-rich schemes, parameterizations will remain essential, particularly for sub-grid cloud processes, ocean mixing and biogeochemical cycles. Enhanced cloud and aerosol microphysics are expected to reduce radiative forcing biases and feedback uncertainties, while improved convective schemes could address persistent bias in thermodynamic processes, precipitation, and surface temperature. Land surface models are also expected to incorporate more realistic vegetation dynamics, improved soil moisture and hydrology schemes, and vertical soil layers, leading to improved land-atmosphere coupling and a better representation of the water-carbon-climate nexus. However, persistent biases in key modes of variability such as ENSO and AMOC are unlikely to be fully resolved, as they emerge from coupled interactions across scales rather than a single model component.

Emerging hybrid approaches that integrate machine learning into parameterizations are expected to reduce uncertainty and improve process representation (Yuval and O'Gorman, 2020; Mansfield and Sheshadri, 2024; Christensen et al., 2024), although issues related to generalisability, physical consistency and interpretability remain active areas of research. These developments will be complemented by strengthened benchmarking and evaluation frameworks, like the Rapid Evaluation Framework (REF) (Hoffman et al., 2025) and standardised diagnostics such as CLIVAR ENSO (Planton et al., 2021a), enabling systematic assessment of model performance across generations.

We also anticipate substantial progress in cryosphere and Earth system coupling. Improved sea dynamics and thermodynamics, together with more realistic ice-sheet-ocean interactions, are expected to reduce biases in sea ice area and freshwater fluxes (Trivedi et al., 2025). Ice-sheet-driven freshwater fluxes can influence ocean circulation and carbon uptake, while associated changes in high-latitude hydrology and permafrost stability may affect methane emissions. As a result, improved representation of permafrost and high-latitude carbon processes is expected to be an important component of CMIP7 (Burke et al., 2017; Turetsky et al., 2020; Park et al., 2023).

CMIP7 will also feature increased emphasis on CO_2-emissions-driven simulations. The introduction of experiments relevant to ZEC within community MIPs like ZECMIP (Jones et al., 2019), CDRMIP (Keller et al., 2018), flat10-MIP (Sanderson et al., 2024, 2025) and TIPMIP (Winkelmann et al., 2025), will enable improved assessment of slow feedbacks, irreversibility, and tipping-point risks. Improved representation of processes governing the TCR and the TCRE is expected to deliver more robust, policy-relevant estimates of climate sensitivity, thereby enabling more policy-relevant constraints on climate sensitivity and improving assessment of threshold behaviour and tipping-point risks.

Despite these advances, it remains unlikely that CMIP7 will '*eliminate*' key structural biases. For both model users and developers, this underscores the continued need for targeted experiments like large ensemble approaches and model physics improvements, as well as process-based evaluation frameworks in the near-future. A shift toward emissions-driven mode in CMIP7 will introduce an additional source of uncertainty arising from carbon-climate feedbacks, including variability in land and ocean carbon uptake. As a result, inter-model spread in temperature and regional climate responses may increase, or become more difficult to attribute to specific physical processes. For model users, this implies a need to explicitly separate uncertainty arising

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

from physical climate processes and that associated with the carbon cycle feedbacks. Large ensembles, reduced-complexity models, carbon-cycle diagnostics, and a systematic comparison between concentration-driven and emissions-driven experiments will be important for isolating mechanisms and ensuring robust interpretation of projected changes.

This review synthesises current understanding of systematic biases across the Earth system, emphasising their interconnected nature and the ways in which they propagate across components. Given the widespread use of CMIP data by a diverse community from students to modelling experts to end users, we provide a high-level assessment of biases in a set of critical diagnostics identified by the CMIP community. We find that reducing systematic biases requires advances in the representation of physical processes, the development, and application of higher-resolution modelling frameworks, and targeted experiments designed to disentangle structural, parametric, and forcing-related sources of uncertainty. Many of these priorities are expected to be addressed in CMIP7. A key strength of the CMIP multimodel framework is the ability to evaluate multiple models against common reference datasets, and ongoing efforts such as the Rapid Evaluation Framework will further support the standardisation and robustness of bias assessment across models.

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Appendix A: Original Survey Questions

Source: WCRP Coupled Model Intercomparison Project (CMIP) Model Benchmarking Task Team and International Project Office

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

CMIP Model Benchmarking

Rapid Evaluation Framework: Diagnostics and reference data

The CMIP Model Benchmarking Task Team are seeking community help to identify a small number of priority diagnostics and their reference datasets that will be incorporated into the design of the Rapid Evaluation Framework (REF) for the AR7 Fast Track.

Find out more about the AR7 Fast Track here: [CMIP Phase 7 (CMIP7) - Coupled Model Intercomparison Project (wcrp-cmip.org)]

This initial focus is on diagnostics with a temporal resolution of monthly means with already have commonly used, open, observation reference data. The intention is to have capability for higher temporal resolutions and we do ask in this survey for your suggestions. We also ask for those aware of new/forthcoming observation datasets, that could be relevant, to identity them.

For full details of the proposed REF, including a 10minute explanatory video see: https://wcrp-cmip.org/ref-announcement/

This survey closes at 7 a.m. UTC on 19 September (slightly extended from the published deadline)

About you

Which of these communities do you belong to? *

Select all that apply

- [x] Diagnostic product user
- [x] Scientific (academic/government scientific institution)
- [x] Modelling centre
- [x] MIP member
- [x] Data provider - observations
- [x] Data infrastructure (e.g.node, ESGF, cloud platform)
- [x] Other

Please list any other communities you belong to here.

Do you have an interest in which diagnostics are included in the essential set for the Rapid Evaluation Framework prototype in time for AR7 Fast Track? *

- (•) Yes
- ( ) No

Clear Selection

Diagnostic themes of interest

- [x] Atmosphere
- [x] Earth System
- [x] Impacts and Adaptation
- [x] Land and land ice
- [x] Oceans and sea ice

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

We are preparing a community owned roadmap for the REF and would like to include initial suggestions from the community. Please suggest higher resolution diagnostics you would like the REF to include and any suggestions for the observation dataset(s).

The reference dataset(s) can be forthcoming

e.g. *Temperature extreme event diagnostic, TX90, daily data - observation suggestion: reanalysis ERA 5*

Observation community

The model benchmarking task team have devised a list of essential diagnostics and reference datasets, organised thematically to match the CMIP Data Request themes. Please take a look at this list before answering the questions in this section: https://airtable.com/applbQctZtl09L2Ga/shrzOD0Hif0PY6XJI

For your datasets relevant to the AR7 Fast Track diagnostics, are they already in obs4MIPs? If not do you plan on submitting by March 2025 (the earlier the better!) so that they are ready in time for prototype testing?

Find out more about obs4MIPs here: [Obs4MIPs in WCRP | Project home (pcmdi.github.io)]()

- ○ Yes, already in obs4MIPs
- ○ Some are already in, the rest will submit by March 2025
- ● Not yet, but will submit by March 2025
- ○ Some are already in, considering submission of others by March 2025
- ○ Not yet, considering submission by March 2025
- ○ No

Clear Selection

Which of the following best describe the datasets you are involved in, relevant to the AR7 Fast Track diagnostics?

Select all that apply

- ☐ Satellite (single)
- ☐ Satellite (merged)
- ☐ Blended Satellite/in situ
- ☐ In-situ (gridded)

Which category best reflects your organisation?

- ○ Space agency
- ○ Academic institution
- ○ Government department
- ○ Meteorological agency/service
- ● Other

Clear Selection

Please indicate alternative category that applies to your organisation

MIP community

Which MIP(s) are you involved with?

Please add all that apply. If you cannot find your MIP here, then it has not yet been registered for CMIP7. Please visit [CMIP Phase 7 (CMIP7) - Coupled Model Intercomparison Project (wcrp-cmip.org)]()

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Are there other non AR7 Fast Track simulations that you think should be evaluated by the REF?

If yes, please set these out.

For your non AR7 Fast Track simulation inclusion suggestions, what metrics would need to be modified to evaluate the data from those simulations, if any?

Feedback on essential diagnostics of the AR7 Fast Track

The model benchmarking task team have devised a list of essential diagnostics, organised thematically to match the CMIP Data Request themes. You can see the full list of diagnostics here: https://airtable.com/applbQctZtl09L2Ga/shrzOD0Hif0PY6XJI

The task team has suggested up to seven metrics per theme. They would appreciate your feedback on the selection of these diagnostics.

For those themes you indicated interest in earlier in the form, you are kindly asked to:

- select up to 5 diagnostics for each theme you feel are important to include and rank these
- indicate your preference for the metric type(s) of each diagnostic you rank.
- check if the reference data set(s) suggested for the different diagnostics would be suitable for the respective diagnostic. If you have a suggestion for an alternative data set please add it in the comment field.

Is there any diagnostic missing that is essential for your requirements at a monthly timescale? *

- (•) Yes
- ( ) No

Clear Selection

Please share your suggestions for essential diagnostic(s)

Please add the name of each missing diagnostic and temporal frequency.

*Please also include the reference data to be used for the diagnostic. Note that we are only looking for diagnostics using a **monthly resolution** for the AR7 Fast track.*

Atmosphere

You can select up to 5 diagnostics. For each of your preferred diagnostics, please rank their importance for inclusion, 1 being highest, 5 being lowest. Only rank those diagnostics that you deem to be important.

Atmosphere - Rank 1	Metric suggestions for Rank 1 diagnostic -can select multiple
+ Add record	+

Atmosphere Rank 2	Metric suggestions for Rank 2
+ Add record	+

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

+ Add record

+

Atmosphere Rank 5

+ Add record

Metric suggestions for Rank 5

+

Please also check if the reference data set(s) suggested for the different diagnostics would be suitable for the respective diagnostics. If you have a comment/suggestion for any alternative data sets please indicate

The table showing suggested reference data sets is here: https://airtable.com/applbQctZtl09L2Ga/shrzOD0Hif0PY6XJI

Make sure you indicate the diagnostic ID they relate to e.g. ID 2.3 Runoff -I disagree with CLASS because.... and would suggest [x] is used instead.

Land and land ice

You can select up to 5 diagnostics. For each of your preferred diagnostics, please rank their importance for inclusion, 1 being highest, 5 being lowest. Only rank those diagnostics that you deem to be important.

Land- Rank 1

+ Add record

Metric suggestions for Rank 1 - can select multiple

+

Land Rank 2

+ Add record

Metric suggestions for Rank 2

+

Land Rank 3

+ Add record

Metric suggestions for Rank 3

+

Land Rank 4

+ Add record

Metric suggestions for Rank 4

+

Land Rank 5

+ Add record

Metric suggestions for Rank 5

+

Please also check if the reference data set(s) suggested for the different diagnostics would be suitable for the respective diagnostics. If you have a comment/suggestion for any alternative data sets please indicate

The table showing suggested reference data sets is here: https://airtable.com/applbQctZtl09L2Ga/shrzOD0Hif0PY6XJI

Make sure you indicate the diagnostic ID they relate to e.g. ID 2.3 Runoff -I disagree with CLASS because.... and would suggest [x] is used instead.

Earth System

You can select up to 5 diagnostics. For each of your preferred diagnostics, please rank their importance for inclusion, 1 being highest, 5 being lowest. Only rank those diagnostics that you deem to be important.

Earth Systems - Rank 1

+ Add record

Metric suggestions for Rank 1 - can select multiple

+

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

+ Add record

+

Earth System Rank 4

+ Add record

Metric suggestions for Rank 4

+

Earth System Rank 5

+ Add record

Metric suggestions for Rank 5

+

Please also check if the reference data set(s) suggested for the different diagnostics would be suitable for the respective diagnostics. If you have a comment/suggestion for any alternative data sets please indicate

The table showing suggested reference data sets is here: https://airtable.com/applbQctZtl09L2Ga/shrzOD0Hif0PY6XJI

Make sure you indicate the diagnostic ID they relate to e.g. ID 2.3 Runoff -I disagree with CLASS because.... and would suggest [x] is used instead.

Impacts and Adaptation

You can select up to 3 diagnostics. For each of your preferred diagnostics, please rank their importance for inclusion, 1 being highest, 3 being lowest. Only rank those diagnostics that you deem to be important.

Impacts - Rank 1

+ Add record

Metric suggestions for Rank 1 - can select multiple

+

Impacts Rank 2

+ Add record

Metric suggestions for Rank 2

+

Impacts Rank 3

+ Add record

Metric suggestions for Rank 3

+

Please also check if the reference data set(s) suggested for the different diagnostics would be suitable for the respective diagnostics. If you have a comment/suggestion for any alternative data sets please indicate

The table showing suggested reference data sets is here: https://airtable.com/applbQctZtl09L2Ga/shrzOD0Hif0PY6XJI

Make sure you indicate the diagnostic ID they relate to e.g. ID 2.3 Runoff -I disagree with CLASS because.... and would suggest [x] is used instead.

Oceans and sea ice

You can select up to 5 diagnostics. For each of your preferred diagnostics, please rank their importance for inclusion, 1 being highest, 5 being lowest. Only rank those diagnostics that you deem to be important.

Oceans- Rank 1

+ Add record

Metric suggestions for Rank 1 - can select multiple

+

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

+ Add record

+

Oceans Rank 4

+ Add record

Metric suggestions for Rank 4

+

Oceans Rank 5

+ Add record

Metric suggestions for Rank 5

+

Please also check if the reference data set(s) suggested for the different diagnostics would be suitable for the respective diagnostics. If you have a comment/suggestion for any alternative data sets please indicate

The table showing suggested reference data sets is here: https://airtable.com/applbQctZtl09L2Ga/shrzOD0Hif0PY6XJI

Make sure you indicate the diagnostic ID they relate to e.g. ID 2.3 Runoff -I disagree with CLASS because.... and would suggest [x] is used instead.

Rapid Evaluation Framework

How do you feel about the development of this community Rapid Evaluation Framework? *

- 5 thumbs up = Very happy
- 4 thumbs up = Happy
- 3 thumbs up = OK
- 2 thumbs up = Not happy
- 1 thumb up = Have no opinion

Is there anything else you'd like to share with us regarding the community Rapid Evaluation Framework?

Do not submit passwords through this form. Report malicious form

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Data availability. Anonymised survey data and code to produce figures can be found in the Zenodo database: https://doi.org/10.5281/zenodo.20927998 (Brown et al., 2026)

Author contributions. Project leads and project design: SB, FB, BP, PB; Survey analysis: AS, FB, AR; Literature review: FB, SB, BP, PB, AS, MSG, RV, SAK, SK, JCB, JC, MD, FF, ML, XSL, YL, AP, SS, SwB; Significant writing contribution: FB, SB, BP, AS, MSG, RV, SAK, SK, SwB; Other writing contribution: JCB, JC, MD, FF, ML, XSL, YL, AP, SS, GS, AL; Editing: SB, BP, PB, FB, AS; Figures: AS, SB, FB, BP.

Writing contributions by Section: Sect. 1: BP; Sect. 2: FB; Sect. 3.1.1: BP, GS; Sect. 3.1.2: SB, SS; Sect. 3.1.3: MD, YL; Sect. 3.1.4: AS; Sect. 3.2.1: SK, SAK; Sect. 3.2.2: RV, FB; Sect. 3.3.1: XSL, AP; Sect. 3.3.2: JCB, MSG; Sect. 3.3.3: AS, JC, SwB; Sect. 3.4.1: ML, RV; Sect. 3.4.2: RV, AS; Sect. 3.5: MSG, SK, SwB; Sect. 4.1: SK, FB; Sect. 4.2: FB, JC; Sect. 5: SK, FB, AS

Competing interests. The authors declare that no competing interests are present.

Acknowledgements. The authors gratefully acknowledge the contributions of the graphic designer Miranta Kouvari from Science Graphic Design to Figures 3-6. We also thank the CMIP-IPO, the Fresh Eyes on CMIP Project leaders, as well as the internal reviewers at CMIP. We finally acknowledge the members of the CMIP community for their feedback through the CMIP REF survey which formed the basis for the analysis presented in this study. SwB acknowledges support from the ATLAS2 project (PID2024-162703OB-I00) funded by MCIN/AEI/10.13039/501100011033 and by ERDF/EU. A.R.R's portion of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004).

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

References

Abatzoglou, J. T. and Kolden, C. A.: Relationships between climate and macroscale area burned in the western United States, International Journal of Wildland Fire, 22, 1003–1020, https://doi.org/10.1071/WF13019, 2013.

Abdelmoaty, H. M., Papalexiou, S. M., Rajulapati, C. R., and AghaKouchak, A.: Biases Beyond the Mean in CMIP6 Extreme Precipitation: A Global Investigation, Earth's Future, 9, e2021EF002 196, https://doi.org/10.1029/2021EF002196, 2021.

Ahn, M.-S., Kim, D., Kang, D., Lee, J., Sperber, K. R., Gleckler, P. J., Jiang, X., Ham, Y.-G., and Kim, H.: MJO Propagation Across the Maritime Continent: Are CMIP6 Models Better Than CMIP5 Models?, Geophysical Research Letters, 47, e2020GL087 250, https://doi.org/10.1029/2020GL087250, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020GL087250, 2020.

Al-Yaari, A., Zhao, Y., Cheruy, F., and Thiery, W.: Heatwave Characteristics in the Recent Climate and at Different Global Warming Levels: A Multimodel Analysis at the Global Scale, Earth's Future, 11, https://doi.org/10.1029/2022EF003301, 2023.

Almazroui, M., Ashfaq, M., Islam, M. N., Rashid, I. U., Kamil, S., Abid, M. A., O'Brien, E., Ismail, M., Reboita, M. S., Sörensson, A. A., Arias, P. A., Alves, L. M., Tippett, M. K., Saeed, S., Haarsma, R., Doblas-Reyes, F. J., Saeed, F., Kucharski, F., Nadeem, I., Silva-Vidal, Y., Rivera, J. A., Ehsan, M. A., Martínez-Castro, D., Muñoz, G., Ali, M. A., Coppola, E., and Sylla, M. B.: Assessment of CMIP6 Performance and Projected Temperature and Precipitation Changes Over South America, Earth Systems and Environment, 5, 155–183, https://doi.org/10.1007/s41748-021-00233-6, 2021.

Anav, A., Friedlingstein, P., Kidston, M., Bopp, L., Ciais, P., Cox, P., Jones, C., Jung, M., Myneni, R., and Zhu, Z.: Evaluating the Land and Ocean Components of the Global Carbon Cycle in the CMIP5 Earth System Models, Journal of Climate, 26, 6801–6843, https://doi.org/10.1175/JCLI-D-12-00417.1, 2013.

Andela, N., Morton, D. C., Giglio, L., Chen, Y., Van Der Werf, G. R., Kasibhatla, P. S., DeFries, R. S., Collatz, G. J., Hantson, S., Kloster, S., Bachelet, D., Forrest, M., Lasslop, G., Li, F., Mangeon, S., Melton, J. R., Yue, C., and Randerson, J. T.: A human-driven decline in global burned area, Science, 356, 1356–1362, https://doi.org/10.1126/science.aal4108, 2017.

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., and Tang, Y.: Forcings, feedbacks, and climate sensitivity in HadGEM3-GC3.1 and UKESM1, Journal of Advances in Modeling Earth Systems, 11, 4377–4394, https://doi.org/https://doi.org/10.1029/2019MS001866, tex.eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019MS001866, 2019.

Anstey, J. A. and Shepherd, T. G.: High-latitude influence of the quasi-biennial oscillation, Quarterly Journal of the Royal Meteorological Society, 140, 1–21, https://doi.org/10.1002/qj.2132, _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.2132, 2014.

Armour, K. C., Bitz, C. M., and Roe, G. H.: Time-varying climate sensitivity from regional feedbacks, Journal of Climate, 26, 4518 – 4534, https://doi.org/10.1175/JCLI-D-12-00544.1, 2013.

Armour, K. C., Proistosescu, C., Dong, Y., Hahn, L. C., Blanchard-Wrigglesworth, E., Pauling, A. G., Jnglin Wills, R. C., Andrews, T., Stuecker, M. F., Po-Chedley, S., Mitevski, I., Forster, P. M., and Gregory, J. M.: Sea-surface temperature pattern effects have slowed global warming and biased warming-based constraints on climate sensitivity, Proceedings of the National Academy of Sciences, 121, e2312093 121, https://doi.org/10.1073/pnas.2312093121, 2024.

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,

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

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, https://doi.org/10.5194/bg-17-4173-2020, 2020.

Baker, J. A., Bell, M. J., Jackson, L. C., Renshaw, R., Vallis, G. K., Watson, A. J., and Wood, R. A.: Overturning Pathways Control AMOC Weakening in CMIP6 Models, Geophysical Research Letters, 50, e2023GL103 381, https://doi.org/10.1029/2023GL103381, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2023GL103381, 2023.

Balch, J. K., Bradley, B. A., Abatzoglou, J. T., Nagy, R. C., Fusco, E. J., and Mahood, A. L.: Human-started wildfires expand the fire niche across the United States, Proceedings of the National Academy of Sciences, 114, 2946–2951, https://doi.org/10.1073/pnas.1617394114, 2017.

Baldwin, M. P., Gray, L. J., Dunkerton, T. J., Hamilton, K., Haynes, P. H., Randel, W. J., Holton, J. R., Alexander, M. J., Hirota, I., Horinouchi, T., Jones, D. B. A., Kinnersley, J. S., Marquardt, C., Sato, K., and Takahashi, M.: The quasi-biennial oscillation, Reviews of Geophysics, 39, 179–229, https://doi.org/10.1029/1999RG000073, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/1999RG000073, 2001.

Bayr, T., Domeisen, D. I. V., and Wengel, C.: The effect of the equatorial Pacific cold SST bias on simulated ENSO teleconnections to the North Pacific and California, Climate Dynamics, 53, 3771–3789, https://doi.org/10.1007/s00382-019-04746-9, 2019.

Bayr, T., Lübbecke, J. F., Vialard, J., and Latif, M.: Equatorial Pacific Cold Tongue Bias Degrades Simulation of ENSO Asymmetry due to Underestimation of Strong Eastern Pacific El Niños, Journal of Climate, 37, 6167–6182, https://doi.org/10.1175/JCLI-D-24-0071.1, 2024.

Beadling, R. L., Russell, J. L., Stouffer, R. J., Mazloff, M., Talley, L. D., Goodman, P. J., Sallée, J.-B., Hewitt, H. T., Hyder, P., and Pandde, A.: Representation of Southern Ocean properties across coupled model intercomparison project generations: CMIP3 to CMIP6, Journal of Climate, 33, 6555–6581, 2020.

Bedia, J., Herrera, S., Gutiérrez, J. M., Benali, A., Brands, S., Mota, B., and Moreno, J. M.: Global patterns in the sensitivity of burned area to fire-weather: Implications for climate change, Agricultural and Forest Meteorology, 214-215, 369–379, https://doi.org/10.1016/j.agrformet.2015.09.002, 2015.

Bellenger, H., Guilyardi, E., Leloup, J., Lengaigne, M., and Vialard, J.: ENSO representation in climate models: from CMIP3 to CMIP5, Climate Dynamics, 42, 1999–2018, https://doi.org/10.1007/s00382-013-1783-z, 2014.

Beniche, M., Vialard, J., Lengaigne, M., and Hall, N. M. J.: How Well Do AMIP6 and CMIP6 Reproduce the Specific Extreme El Niño Teleconnections to North America?, Journal of Geophysical Research: Atmospheres, 130, e2024JD041 740, https://doi.org/10.1029/2024JD041740, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2024JD041740, 2025.

Berg, A. and Sheffield, J.: Historic and Projected Changes in Coupling Between Soil Moisture and Evapotranspiration (ET) in CMIP5 Models Confounded by the Role of Different ET Components, Journal of Geophysical Research: Atmospheres, 124, 5791–5806, https://doi.org/10.1029/2018JD029807, 2019.

Biastoch, A., Schwarzkopf, F. U., Getzlaff, K., Rühs, S., Martin, T., Scheinert, M., Schulzki, T., Handmann, P., Hummels, R., and Böning, C. W.: Regional imprints of changes in the Atlantic Meridional Overturning Circulation in the eddy-rich ocean model VIKING20X, Ocean Science, 17, 1177–1211, https://doi.org/10.5194/os-17-1177-2021, 2021.

Bock, L., Lauer, A., Schlund, M., Barreiro, M., Bellouin, N., Jones, C., Meehl, G. A., Predoi, V., Roberts, M. J., and Eyring, V.: Quantifying Progress Across Different CMIP Phases With the ESMValTool, Journal of Geophysical Research: Atmospheres, 125, e2019JD032 321, https://doi.org/10.1029/2019JD032321, 2020.

Bodas-Salcedo, A., Webb, M., Brooks, M., Ringer, M., Williams, K., Milton, S., and Wilson, D.: Evaluating cloud systems in the Met Office global forecast model using simulated CloudSat radar reflectivities, Journal of Geophysical Research: Atmospheres, 113, 2008.

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Bodas-Salcedo, A., Williams, K. D., Ringer, M. A., Beau, I., Cole, J. N., Dufresne, J.-L., Koshiro, T., Stevens, B., Wang, Z., and Yokohata, T.: Origins of the solar radiation biases over the Southern Ocean in CFMIP2 models, Journal of Climate, 27, 41–56, 2014.

Bodas-Salcedo, A., Mulcahy, J. P., Andrews, T., Williams, K. D., Ringer, M. A., Field, P. R., and Elsaesser, G. S.: Strong dependence of atmospheric feedbacks on mixed-phase microphysics and aerosol-cloud interactions in HadGEM3, Journal of Advances in Modeling Earth Systems, 11, 1735–1758, https://doi.org/https://doi.org/10.1029/2019MS001688, tex.eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019MS001688, 2019.

Bonan, D. B., Thompson, A. F., Schneider, T., Zanna, L., Armour, K. C., and Sun, S.: Observational constraints imply limited future Atlantic meridional overturning circulation weakening, Nature Geoscience, 18, 479–487, https://doi.org/10.1038/s41561-025-01709-0, 2025.

Bony, S., Colman, R., Kattsov, V. M., Allan, R. P., Bretherton, C. S., Dufresne, J.-L., Hall, A., Hallegatte, S., Holland, M. M., Ingram, W., Randall, D. A., Soden, B. J., Tselioudis, G., and Webb, M. J.: How Well Do We Understand and Evaluate Climate Change Feedback Processes?, Journal of Climate, 19, 3445–3482, https://doi.org/10.1175/JCLI3819.1, 2006.

Borato, L., Härter Fetter Filho, A. F., Gomes Da Silva, P., Mendez, F. J., and Da Fontoura Klein, A. H.: Evaluation of CMIP5 Models Based on Weather Types Applied to the South Atlantic Ocean, International Journal of Climatology, 44, 5580–5595, https://doi.org/10.1002/joc.8653, 2024.

Borowiak, A., King, A. D., Brown, J. R., Jones, C. D., and Grose, M.: Climate Stabilisation Under Net Zero CO2 Emissions, Earth's Future, 13, e2024EF005 678, https://doi.org/10.1029/2024EF005678, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2024EF005678, 2025a.

Borowiak, A., King, A. D., Brown, J. R., and Ziehn, T.: Revised estimates of temperature changes under net zero CO2 emissions, npj Climate and Atmospheric Science, 8, 275, https://doi.org/10.1038/s41612-025-01148-z, 2025b.

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.: 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, 2013.

Brands, S.: A circulation-based performance atlas of the CMIP5 and 6 models for regional climate studies in the Northern Hemisphere mid-to-high latitudes, Geoscientific Model Development, 15, 1375–1411, https://doi.org/10.5194/gmd-15-1375-2022, 2022a.

Brands, S.: Common Error Patterns in the Regional Atmospheric Circulation Simulated by the CMIP Multi-Model Ensemble, Geophysical Research Letters, 49, e2022GL101 446, https://doi.org/10.1029/2022GL101446, 2022b.

Brands, S., Fernández-Granja, J. A., Bedia, J., Casanueva, A., and Fernández, J.: A Global Climate Model Performance Atlas for the Southern Hemisphere Extratropics Based on Regional Atmospheric Circulation Patterns, Geophysical Research Letters, 50, e2023GL103 531, https://doi.org/10.1029/2023GL103531, 2023.

Brown, F., Bischof, S., and Pearson, B.: Dataset for: Fresh Eyes on CMIP Model Biases: Diagnosing Biases Across CMIP Generations and Implications for CMIP, https://doi.org/10.5281/zenodo.20927998, 2026.

Brown, J. R., Brierley, C. M., An, S.-I., Guarino, M.-V., Stevenson, S., Williams, C. J. R., Zhang, Q., Zhao, A., Abe-Ouchi, A., Braconnot, P., Brady, E. C., Chandan, D., D'Agostino, R., Guo, C., LeGrande, A. N., Lohmann, G., Morozova, P. A., Ohgaito, R., O'ishi, R., Otto-Bliesner, B. L., Peltier, W. R., Shi, X., Sime, L., Volodin, E. M., Zhang, Z., and Zheng, W.: Comparison of past and future simulations of ENSO in CMIP5/PMIP3 and CMIP6/PMIP4 models, Climate of the Past, 16, 1777–1805, https://doi.org/10.5194/cp-16-1777-2020, 2020.

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Brown, P. J., McDonagh, E. L., Sanders, R., Watson, A. J., Wanninkhof, R., King, B. A., Smeed, D. A., Baringer, M. O., Meinen, C. S., Schuster, U., Yool, A., and Messias, M.-J.: Circulation-driven variability of Atlantic anthropogenic carbon transports and uptake, Nature Geoscience, 14, 571–577, https://doi.org/10.1038/s41561-021-00774-5, 2021.

Bryden, H., Beunk, J., Drijfhout, S., Hazeleger, W., and Mecking, J.: Comparing observed and modelled components of the Atlantic Meridional Overturning Circulation at 26° N, Ocean Science, 20, 589–599, https://doi.org/10.5194/os-20-589-2024, 2024.

Burke, E. J., Ekici, A., Huang, Y., Chadburn, S. E., Huntingford, C., Ciais, P., Friedlingstein, P., Peng, S., and Krinner, G.: Quantifying uncertainties of permafrost carbon–climate feedbacks, Biogeosciences, 14, 3051–3066, https://doi.org/10.5194/bg-14-3051-2017, 2017.

Burke, E. J., Zhang, Y., and Krinner, G.: Evaluating permafrost physics in the Coupled Model Intercomparison Project 6 (CMIP6) models and their sensitivity to climate change, The Cryosphere, 14, 3155–3174, https://doi.org/10.5194/tc-14-3155-2020, 2020.

Butchart, N., Andrews, M. B., and Jones, C. D.: QBO Phase Synchronization in CMIP6 Historical Simulations Attributed to Ozone Forcing, Geophysical Research Letters, 50, e2023GL104 401, https://doi.org/10.1029/2023GL104401, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2023GL104401, 2023.

Calvo-Sancho, C., Díaz-Fernández, J., González-Alemán, J. J., Halifa-Marín, A., Miglietta, M. M., Azorin-Molina, C., Prein, A. F., Montoro-Mendoza, A., Bolgiani, P., Morata, A., and Martín, M. L.: Human-induced climate change amplification on storm dynamics in Valencia's 2024 catastrophic flash flood, Nature Communications, 17, 1492, https://doi.org/10.1038/s41467-026-68929-9, 2026.

Canadell, J. G., Monteiro, P. M., Costa, M. H., Cunha, L. C. D., Cox, P. M., Eliseev, A. V., Henson, S., Ishii, M., Jaccard, S., Koven, C., Lohila, A., Patra, P. K., Piao, S., Syampungani, S., Zaehle, S., Zickfeld, K., Alexandrov, G. A., Bala, G., Bopp, L., Boysen, L., Cao, L., Chandra, N., Ciais, P., Denisov, S. N., Dentener, F. J., Douville, H., Fay, A., Forster, P., Fox-Kemper, B., Friedlingstein, P., Fu, W., Fuss, S., Garcon, V., Gier, B., Gillett, N. P., Gregor, L., Haustein, K., Haverd, V., He, J., Hewitt, H. T., Hoffman, F. M., Ilyina, T., Jackson, R., Jones, C., Keller, D. P., Kwiatkowski, L., Lamboll, R. D., Lan, X., Laufkötter, C., Le Quéré, C., Lenton, A., Lewis, J., Liddicoat, S., Lorenzoni, L., Lovenduski, N., Macdougall, A. H., Mathesius, S., Matthews, D. H., Meinshausen, M., Mokhov, I. I., Naik, V., Nicholls, Z. R. J., Nurhati, I. S., O'sullivan, M., Peters, G., Pongratz, J., Poulter, B., Sallée, J.-B., Saunois, M., Schuur, E. A., I.Seneviratne, S., Stavert, A., Suntharalingam, P., Tachiiri, K., Terhaar, J., Thompson, R., Tian, H., Turnbull, J., Vicente-Serrano, S. M., Wang, X., Wanninkhof, R. H., Williamson, P., Brovkin, V., Feely, R. A., and Lebehot, A. D.: Global Carbon and other Biogeochemical Cycles and Feedbacks, in: IPCC AR6 WGI, Final Government Distribution, p. chapter 5, https://hal.science/hal-03336145, 2021.

Cannon, A. J.: Reductions in daily continental-scale atmospheric circulation biases between generations of global climate models: CMIP5 to CMIP6, Environmental Research Letters, 15, 064 006, https://doi.org/10.1088/1748-9326/ab7e4f, 2020.

Capotondi, A., McGregor, S., McPhaden, M. J., Cravatte, S., Holbrook, N. J., Imada, Y., Sanchez, S. C., Sprintall, J., Stuecker, M. F., Ummenhofer, C. C., Zeller, M., Farneti, R., Graffino, G., Hu, S., Karnauskas, K. B., Kosaka, Y., Kucharski, F., Mayer, M., Qiu, B., Santoso, A., Taschetto, A. S., Wang, F., Zhang, X., Holmes, R. M., Luo, J.-J., Maher, N., Martinez-Villalobos, C., Meehl, G. A., Naha, R., Schneider, N., Stevenson, S., Sullivan, A., van Rensch, P., and Xu, T.: Mechanisms of tropical Pacific decadal variability, Nature Reviews Earth & Environment, 4, 754–769, https://doi.org/10.1038/s43017-023-00486-x, 2023.

Cardil, A., Eastaugh, C. S., and Molina, D. M.: Extreme temperature conditions and wildland fires in Spain, Theoretical and Applied Climatology, 122, 219–228, https://doi.org/10.1007/s00704-014-1295-8, 2015.

Casagrande, F., Stachelski, L., and De Souza, R. B.: Assessment of Antarctic sea ice area and concentration in Coupled Model Intercomparison Project Phase 5 and Phase 6 models, International Journal of Climatology, 43, 1314–1332, https://doi.org/10.1002/joc.7916, 2023.

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Cassidy, L. J., King, A. D., Brown, J., MacDougall, A. H., Ziehn, T., Min, S.-K., and Jones, C. D.: Regional temperature extremes and vulnerability under net zero CO2 emissions, Environmental Research Letters, https://iopscience.iop.org/article/10.1088/1748-9326/ad114a/meta, 2023.

Cassidy, L. J., King, A. D., Brown, J. R., Ziehn, T., and Borowiak, A.: Evolution of Heat Extremes Under Net-Zero CO2 Emissions, Earth's Future, 13, e2024EF005 399, https://doi.org/10.1029/2024EF005399, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2024EF005399, 2025.

Cesana, G. V., Khadir, T., Chepfer, H., and Chiriaco, M.: Southern Ocean solar reflection biases in CMIP6 models linked to cloud phase and vertical structure representations, Geophysical Research Letters, 49, e2022GL099 777, 2022.

Chadburn, S. E., Krinner, G., Porada, P., Bartsch, A., Beer, C., Belelli Marchesini, L., Boike, J., Ekici, A., Elberling, B., Friborg, T., Hugelius, G., Johansson, M., Kuhry, P., Kutzbach, L., Langer, M., Lund, M., Parmentier, F.-J. W., Peng, S., Van Huissteden, K., Wang, T., Westermann, S., Zhu, D., and Burke, E. J.: Carbon stocks and fluxes in the high latitudes: using site-level data to evaluate Earth system models, Biogeosciences, 14, 5143–5169, https://doi.org/10.5194/bg-14-5143-2017, 2017.

Chen, H.-C. and Jin, F.-F.: Dynamics of ENSO Phase–Locking and Its Biases in Climate Models, Geophysical Research Letters, 49, e2021GL097 603, https://doi.org/10.1029/2021GL097603, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2021GL097603, 2022.

Chen, J., Yang, Y., and Tang, J.: Bias correction of surface air temperature and precipitation in CORDEX East Asia simulation: What should we do when applying bias correction?, Atmospheric Research, 280, https://doi.org/10.1016/j.atmosres.2022.106439, 2022.

Choudhury, B. A., Rajesh, P. V., Zahan, Y., and Goswami, B. N.: Evolution of the Indian summer monsoon rainfall simulations from CMIP3 to CMIP6 models, Climate Dynamics, 58, 2637–2662, https://doi.org/10.1007/s00382-021-06023-0, 2022.

Christensen, H. M., Kouhen, S., Miller, G., and Parthipan, R.: Machine learning for stochastic parametrization, Environmental Data Science, 3, e38, https://doi.org/10.1017/eds.2024.45, 2024.

Chung, C., Boschat, G., Taschetto, A., Narsey, S., McGregor, S., Santoso, A., and Delage, F.: Evaluation of seasonal teleconnections to remote drivers of Australian rainfall in CMIP5 and CMIP6 models, Journal of Southern Hemisphere Earth Systems Science, 73, 219–261, https://www.publish.csiro.au/es/ES23002, 2023.

Chung, C. T. Y., Power, S. B., Sullivan, A., and Delage, F.: The role of the South Pacific in modulating Tropical Pacific variability, Scientific Reports, 9, 18 311, https://doi.org/10.1038/s41598-019-52805-2, 2019.

Coburn, J. and Pryor, S. C.: Differential Credibility of Climate Modes in CMIP6, Journal of Climate, 34, 8145–8164, https://doi.org/10.1175/JCLI-D-21-0359.1, 2021.

Collimore, C. C., Martin, D. W., Hitchman, M. H., Huesmann, A., and Waliser, D. E.: On The Relationship between the QBO and Tropical Deep Convection, Journal of Climate, 16, 2552–2568, https://doi.org/10.1175/1520-0442(2003)016<2552:OTRBTQ>2.0.CO;2, 2003.

Cook, B. I., Mankin, J. S., Marvel, K., Williams, A. P., Smerdon, J. E., and Anchukaitis, K. J.: Twenty-First Century Drought Projections in the CMIP6 Forcing Scenarios, Earth's Future, 8, e2019EF001 461, https://doi.org/10.1029/2019EF001461, 2020.

Cox, P. M., Williamson, M. S., Friedlingstein, P., Jones, C. D., Raoult, N., Rogelj, J., and Varney, R. M.: Emergent constraints on carbon budgets as a function of global warming, Nature Communications, 15, 1885, https://doi.org/10.1038/s41467-024-46137-7, 2024.

Davini, P. and D'Andrea, F.: Northern Hemisphere Atmospheric Blocking Representation in Global Climate Models: Twenty Years of Improvements?, Journal of Climate, 29, 8823–8840, https://doi.org/10.1175/JCLI-D-16-0242.1, 2016.

Davini, P. and D'Andrea, F.: From CMIP3 to CMIP6: Northern Hemisphere Atmospheric Blocking Simulation in Present and Future Climate, Journal of Climate, 33, 10 021–10 038, https://doi.org/10.1175/JCLI-D-19-0862.1, 2020.

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

De Luca, P., Jiménez-Esteve, B., Degenhardt, L., Schemm, S., and Pfahl, S.: Enhanced Blocking Frequencies in Very-High Resolution Idealized Climate Model Simulations, Geophysical Research Letters, 51, e2024GL111 016, https://doi.org/10.1029/2024GL111016, 2024.

Delhasse, A., Hanna, E., Kittel, C., and Fettweis, X.: Brief communication: CMIP6 does not suggest any atmospheric blocking increase in summer over Greenland by 2100, International Journal of Climatology, 41, 2589–2596, https://doi.org/10.1002/joc.6977, 2021.

Desai, A. R., Richardson, A. D., Moffat, A. M., Kattge, J., Hollinger, D. Y., Barr, A., Falge, E., Noormets, A., Papale, D., Reichstein, M., and Stauch, V. J.: Cross-site evaluation of eddy covariance GPP and RE decomposition techniques, Agricultural and Forest Meteorology, 148, 821–838, https://doi.org/10.1016/j.agrformet.2007.11.012, 2008.

Devilliers, M., Swingedouw, D., Mignot, J., Deshayes, J., Garric, G., and Ayache, M.: A realistic Greenland ice sheet and surrounding glaciers and ice caps melting in a coupled climate model, Climate Dynamics, 57, 2467–2489, https://doi.org/10.1007/s00382-021-05816-7, 2021.

Devilliers, M., Yang, S., Drews, A., Schmith, T., and Olsen, S. M.: Ocean response to a century of observation-based freshwater forcing around Greenland in EC-Earth3, Climate Dynamics, 62, 4905–4923, https://doi.org/10.1007/s00382-024-07142-0, 2024.

Di Virgilio, G., Evans, J. P., Blake, S. A. P., Armstrong, M., Dowdy, A. J., Sharples, J., and McRae, R.: Climate Change Increases the Potential for Extreme Wildfires, Geophysical Research Letters, 46, 8517–8526, https://doi.org/10.1029/2019GL083699, 2019.

Diamond, R., Sime, L. C., Holmes, C. R., and Schroeder, D.: CMIP6 models rarely simulate Antarctic winter sea-ice anomalies as large as observed in 2023, Geophysical Research Letters, 51, e2024GL109 265, 2024.

Dittus, A. J., Collins, M., Sutton, R., and Hawkins, E.: Reversal of Projected European Summer Precipitation Decline in a Stabilizing Climate, Geophysical Research Letters, 51, e2023GL107 448, https://doi.org/10.1029/2023GL107448, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2023GL107448, 2024.

Dong, J., Lei, F., and Crow, W.: Land transpiration-evaporation partitioning errors responsible for modeled summertime warm bias in the central United States, Nature Communications, 13, https://doi.org/10.1038/s41467-021-27938-6, 2022.

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, https://doi.org/10.1175/JCLI-D-19-1011.1, 2020.

Douville, H. and Cheng, L.: Asymmetric Sea Surface Salinity Response to Global Warming: “Fresh Gets Fresher but Salty Hesitates”, Geophysical Research Letters, 51, e2023GL107 944, https://doi.org/10.1029/2023GL107944, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2023GL107944, 2024.

Dunkl, I., Lovenduski, N., Collalti, A., Arora, V. K., Ilyina, T., and Brovkin, V.: Gross primary productivity and the predictability of CO_2: more uncertainty in what we predict than how well we predict it, Biogeosciences, 20, 3523–3538, https://doi.org/10.5194/bg-20-3523-2023, 2023.

Dunne, J. P., Hewitt, H. T., Arblaster, J. M., Bonou, F., Boucher, O., Cavazos, T., Dingley, B., Durack, P. J., Hassler, B., Juckes, M., Miyakawa, T., Mizielinski, M., Naik, V., Nicholls, Z., O’Rourke, E., Pincus, R., Sanderson, B. M., Simpson, I. R., and Taylor, K. E.: An evolving Coupled Model Intercomparison Project phase 7 (CMIP7) and Fast Track in support of future climate assessment, Geoscientific Model Development, 18, 6671–6700, https://doi.org/10.5194/gmd-18-6671-2025, 2025.

Dutta, R. and Maity, R.: Value addition in coupled model intercomparison project phase 6 over phase 5: global perspectives of precipitation, temperature and soil moisture fields, Acta Geophysica, 70, 1401–1415, https://doi.org/10.1007/s11600-022-00793-9, 2022.

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Eayrs, C., Holland, D., Francis, D., Wagner, T., Kumar, R., and Li, X.: Understanding the Seasonal Cycle of Antarctic Sea Ice Extent in the Context of Longer-Term Variability, Reviews of Geophysics, 57, 1037–1064, https://doi.org/https://doi.org/10.1029/2018RG000631, 2019.

Ebdon, R. A. and Veryard, R. G.: Fluctuations in Equatorial Stratospheric Winds, Nature, 189, 791–793, https://doi.org/10.1038/189791a0, 1961.

Elbaum, E., Garfinkel, C. I., Adam, O., Morin, E., Rostkier-Edelstein, D., and Dayan, U.: Uncertainty in Projected Changes in Precipitation Minus Evaporation: Dominant Role of Dynamic Circulation Changes and Weak Role for Thermodynamic Changes, Geophysical Research Letters, 49, e2022GL097 725, https://doi.org/10.1029/2022GL097725, 2022.

England, M. R., Polvani, L. M., Screen, J., and Chan, A. C.: Minimal Arctic Sea Ice Loss in the Last 20 Years, Consistent With Internal Climate Variability, Geophysical Research Letters, 52, e2025GL116 175, https://doi.org/10.1029/2025GL116175, 2025.

EUMETSAT OSI SAF and EUMETSAT SAF on Ocean and Sea Ice: Sea ice index version 3.0 - multimission, https://doi.org/10.15770/EUM_SAF_OSI_0026, tex.copyright: Free and unrestricted - CC-BY-4.0, 2026.

Exarchou, E., Prodhomme, C., Brodeau, L., Guemas, V., and Doblas-Reyes, F.: Origin of the warm eastern tropical Atlantic SST bias in a climate model, Climate Dynamics, 51, 1819–1840, https://doi.org/10.1007/s00382-017-3984-3, 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, Geoscientific Model Development, 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Fasullo, J. T., Phillips, A. S., and Deser, C.: Evaluation of Leading Modes of Climate Variability in the CMIP Archives, Journal of Climate, 33, 5527–5545, https://doi.org/10.1175/JCLI-D-19-1024.1, 2020.

Fathrio, I., Iizuka, S., Manda, A., Kodama, Y.-M., Ishida, S., Moteki, Q., Yamada, H., and Tachibana, Y.: Assessment of western I ndian O cean SST bias of CMIP5 models, Journal of Geophysical Research: Oceans, 122, 3123–3140, 2017.

Feng, S., Zscheischler, J., Hao, Z., and Bevacqua, E.: Growing human-induced climate change fingerprint in regional weekly fire extremes, npj Climate and Atmospheric Science, 8, 152, https://doi.org/10.1038/s41612-025-01021-z, 2025.

Fernández-Guisuraga, J. M., Suárez-Seoane, S., García-Llamas, P., and Calvo, L.: Vegetation structure parameters determine high burn severity likelihood in different ecosystem types: A case study in a burned Mediterranean landscape, Journal of Environmental Management, 288, 112 462, https://doi.org/10.1016/j.jenvman.2021.112462, 2021.

Ferrett, S., Collins, M., Ren, H.-L., Wu, B., and Zhou, T.: The Role of Tropical Mean-State Biases in Modeled Winter Northern Hemisphere El Niño Teleconnections, Journal of Climate, 33, 4751–4768, https://doi.org/10.1175/JCLI-D-19-0668.1, 2020.

Flannigan, M. D. and Harrington, J. B.: A Study of the Relation of Meteorological Variables to Monthly Provincial Area Burned by Wildfire in Canada (1953–80), Journal of Applied Meteorology, 27, 441–452, https://doi.org/10.1175/1520-0450(1988)027<0441:ASOTRO>2.0.CO;2, 1988.

Forster, P. M., Storelvmo, T., Armour, K., Collins, W., Dufresne, J.-L., Frame, D., Lunt, D., Mauritsen, T., Palmer, M. D., Watanabe, M., Wild, M., and Zhang, H.: The earth's energy budget, climate feedbacks, and climate sensitivity, 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, https://doi.org/10.1017/9781009157896.009, 2021.

Fox-Kemper, B., Hewitt, H. T., Xiao, C., Aðalgeirsdóttir, G., Drijfhout, S. S., Edwards, T. L., Golledge, N. R., Hemer, M., Kopp, R. E., Krinner, G., Mix, A., Notz, D., Nowicki, S., Nurhati, I. S., Ruiz, L., Sallée, J.-B., Slangen, A. B. A., and Yu, Y.: Ocean, cryosphere and sea level change, in: Climate change 2021: The physical science basis, 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., pp. 1211–1362, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/10.1017/9781009157896.011, 2021.

Freund, M. B., Brown, J. R., Henley, B. J., Karoly, D. J., and Brown, J. N.: Warming Patterns Affect El Niño Diversity in CMIP5 and CMIP6 Models, Journal of Climate, 33, 8237–8260, https://doi.org/10.1175/JCLI-D-19-0890.1, 2020.

Freund, M. B., Brown, J. R., Marshall, A. G., Tozer, C. R., Henley, B. J., Risbey, J. S., Ramesh, N., Lieber, R., and Sharmila, S.: Interannual ENSO diversity, transitions, and projected changes in observations and climate models, Environmental Research Letters, 19, 114 005, https://doi.org/10.1088/1748-9326/ad78db, 2024.

Fyfe, J. C., Kharin, V. V., Santer, B. D., Cole, J. N. S., and Gillett, N. P.: Significant impact of forcing uncertainty in a large ensemble of climate model simulations, Proceedings of the National Academy of Sciences, 118, e2016549 118, https://doi.org/10.1073/pnas.2016549118, 2021.

Gaboriau, D. M., Remy, C. C., Girardin, M. P., Asselin, H., Hély, C., Bergeron, Y., and Ali, A. A.: Temperature and fuel availability control fire size/severity in the boreal forest of central Northwest Territories, Canada, Quaternary Science Reviews, 250, 106 697, https://doi.org/10.1016/j.quascirev.2020.106697, 2020.

Gallo, C., Eden, J. M., Dieppois, B., Drobyshev, I., Fulé, P. Z., San-Miguel-Ayanz, J., and Blackett, M.: Evaluation of CMIP6 model performances in simulating fire weather spatiotemporal variability on global and regional scales, Geoscientific Model Development, 16, 3103–3122, https://doi.org/10.5194/gmd-16-3103-2023, 2023.

Geller, M. A., Zhou, T., Shindell, D., Ruedy, R., Aleinov, I., Nazarenko, L., Tausnev, N. L., Kelley, M., Sun, S., Cheng, Y., Field, R. D., and Faluvegi, G.: Modeling the QBO—Improvements resulting from higher-model vertical resolution, Journal of Advances in Modeling Earth Systems, 8, 1092–1105, https://doi.org/10.1002/2016MS000699, 2016.

Gier, B. K., Schlund, M., Friedlingstein, P., Jones, C. D., Jones, C., Zaehle, S., and Eyring, V.: Representation of the terrestrial carbon cycle in CMIP6, Biogeosciences, 21, 5321–5360, https://doi.org/10.5194/bg-21-5321-2024, 2024.

Goddard, L. and Gershunov, A.: Impact of El Niño on Weather and Climate Extremes, in: El Niño Southern Oscillation in a Changing Climate, pp. 361–375, American Geophysical Union (AGU), ISBN 978-1-119-54816-4, https://doi.org/10.1002/9781119548164.ch16, section: 16 _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/9781119548164.ch16, 2020.

Goelzer, H., Nowicki, S., Payne, A., Larour, E., Seroussi, H., Lipscomb, W. H., Gregory, J., Abe-Ouchi, A., Shepherd, A., Simon, E., Agosta, C., Alexander, P., Aschwanden, A., Barthel, A., Calov, R., Chambers, C., Choi, Y., Cuzzone, J., Dumas, C., Edwards, T., Felikson, D., Fettweis, X., Golledge, N. R., Greve, R., Humbert, A., Huybrechts, P., Le Clec'H, S., Lee, V., Leguy, G., Little, C., Lowry, D. P., Morlighem, M., Nias, I., Quiquet, A., Rückamp, M., Schlegel, N.-J., Slater, D. A., Smith, R. S., Straneo, F., Tarasov, L., Van De Wal, R., and Van Den Broeke, M.: The future sea-level contribution of the Greenland ice sheet: a multi-model ensemble study of ISMIP6, The Cryosphere, 14, 3071–3096, https://doi.org/10.5194/tc-14-3071-2020, 2020.

Gollan, G., Bastin, S., and Greatbatch, R. J.: Tropical precipitation influencing boreal winter midlatitude blocking, Atmospheric Science Letters, 20, e900, https://doi.org/10.1002/asl.900, 2019.

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Goosse, H., Allende Contador, S., Bitz, C. M., Blanchard-Wrigglesworth, E., Eayrs, C., Fichefet, T., Himmich, K., Huot, P.-V., Klein, F., and Marchi, S.: Modulation of the seasonal cycle of the Antarctic sea ice extent by sea ice processes and feedbacks with the ocean and the atmosphere, The Cryosphere Discussions, 2022, 1–30, 2022.

Gopika, S., Lengaigne, M., Suresh, I., Izumo, T., Kwatra, S., Neetu, S., and Vialard, J.: Assessing CMIP Models' Ability to Detect Observed Surface Warming Signals Related to Climate Change, Journal of Climate, https://doi.org/10.1175/JCLI-D-24-0102.1, 2024.

Gopika, S., Sadhvi, K., Vialard, J., Danielli, V., Neetu, S., and Lengaigne, M.: Drivers of future Indian Ocean warming and its spatial pattern in CMIP models,, Earth's Future, 13, e2025EF006 112, 2025.

Green, J. K.: The intricacies of vegetation responses to changing moisture conditions, New Phytologist, 244, 2156–2162, https://doi.org/10.1111/nph.20182, _eprint: https://nph.onlinelibrary.wiley.com/doi/pdf/10.1111/nph.20182, 2024.

Guilyardi, E., Capotondi, A., Lengaigne, M., Thual, S., and Wittenberg, A. T.: ENSO Modeling, in: El Niño Southern Oscillation in a Changing Climate, pp. 199–226, American Geophysical Union (AGU), ISBN 978-1-119-54816-4, https://doi.org/10.1002/9781119548164.ch9, section: 9 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781119548164.ch9, 2020.

Guimarães, S. O., Mann, M. E., Rahmstorf, S., Petri, S., Steinman, B. A., Brouillette, D. J., Christiansen, S., and Li, X.: Increased projected changes in quasi-resonant amplification and persistent summer weather extremes in the latest multimodel climate projections, Scientific Reports, 14, 21 991, https://doi.org/10.1038/s41598-024-72787-0, 2024.

Guo, H., Zhan, C., Ning, L., Li, Z., and Hu, S.: Evaluation and comparison of CMIP6 and CMIP5 model performance in simulating the runoff, Theoretical and Applied Climatology, 149, 1451–1470, https://doi.org/https://doi.org/10.1007/s00704-022-04118-0, 2022.

Gurazada, M., McDermid, S., DeFries, R., Davis, K. F., Singh, J., and Singh, D.: El Niño and positive Indian Ocean Dipole conditions simultaneously reduce the production of multiple cereals across India, Environmental Research Letters, 19, 104 059, https://doi.org/10.1088/1748-9326/ad6a6f, 2024.

Harris, S., Nicholls, N., Tapper, N., and Mills, G.: The sensitivity of fire activity to interannual climate variability in Victoria, Australia, Journal of Southern Hemisphere Earth Systems Science, 69, 146–160, https://doi.org/10.1071/ES19008, 2020.

Hawkins, E. and Sutton, R.: The potential to narrow uncertainty in regional climate predictions, Bulletin of the American Meteorological Society, 90, 1095–1108, 2009.

He, L., Zhou, T., and Chen, X.: South Asian summer rainfall from CMIP3 to CMIP6 models: biases and improvements, Climate Dynamics, 61, 1049–1061, https://doi.org/10.1007/s00382-022-06542-4, 2023.

Hernandez, M. and Chen, L.: Future Land Precipitation Changes Over the North American Monsoon Region Using CMIP5 and CMIP6 Simulations, Journal of Geophysical Research: Atmospheres, 127, e2021JD035 911, https://doi.org/10.1029/2021JD035911, 2022.

Heuzé, C.: Antarctic Bottom Water and North Atlantic Deep Water in CMIP6 models, Ocean Science, 17, 59–90, https://doi.org/10.5194/os-17-59-2021, 2021.

Hirschi, J. J., Barnier, B., Böning, C., Biastoch, A., Blaker, A. T., Coward, A., Danilov, S., Drijfhout, S., Getzlaff, K., Griffies, S. M., Hasumi, H., Hewitt, H., Iovino, D., Kawasaki, T., Kiss, A. E., Koldunov, N., Marzocchi, A., Mecking, J. V., Moat, B., Molines, J., Myers, P. G., Penduff, T., Roberts, M., Treguier, A., Sein, D. V., Sidorenko, D., Small, J., Spence, P., Thompson, L., Weijer, W., and Xu, X.: The Atlantic Meridional Overturning Circulation in High-Resolution Models, Journal of Geophysical Research: Oceans, 125, e2019JC015 522, https://doi.org/10.1029/2019JC015522, 2020.

Hobbs, W., Spence, P., Meyer, A., Schroeter, S., Fraser, A. D., Reid, P., Tian, T. R., Wang, Z., Liniger, G., Doddridge, E. W., and Boyd, P. W.: Observational Evidence for a Regime Shift in Summer Antarctic Sea Ice, Journal of Climate, 37, 2263–2275, https://doi.org/10.1175/JCLI-D-23-0479.1, 2024.

Hobbs, W. R., Massom, R., Stammerjohn, S., Reid, P., Williams, G., and Meier, W.: A review of recent changes in Southern Ocean sea ice, their drivers and forcings, Global and Planetary Change, 143, 228–250, https://doi.org/10.1016/j.gloplacha.2016.06.008, 2016.

Hoffman, F. M., Hassler, B., Swaminathan, R., Lewis, J., Andela, B., Collier, N., Hegedűs, D., Lee, J., Pascoe, C., Pflüger, M., Stockhause, M., Ullrich, P., Xu, M., Bock, L., Chun, F., Gier, B. K., Kelley, D. I., Lauer, A., Lenhardt, J., Schlund, M., Sreeush, M. G., Weigel, K., Blockley, E., Beadling, R., Beucher, R., Dugassa, D. D., Lembo, V., Lu, J., Brands, S., Tjiputra, J., Malinina, E., Mederios, B., Scoccimarro, E., Walton, J., Kershaw, P., Marquez, A. L., Roberts, M. J., O'Rourke, E., Dingley, E., Turner, B., Hewitt, H., and Dunne, J. P.: Rapid Evaluation Framework for the CMIP7 Assessment Fast Track, https://doi.org/10.5194/egusphere-2025-2685, 2025.

Holden, Z. A., Swanson, A., Luce, C. H., Jolly, W. M., Maneta, M., Oyler, J. W., Warren, D. A., Parsons, R., and Affleck, D.: Decreasing fire season precipitation increased recent western US forest wildfire activity, Proceedings of the National Academy of Sciences, 115, https://doi.org/10.1073/pnas.1802316115, 2018.

Holmes, C. R., Bracegirdle, T. J., Holland, P. R., Stroeve, J., and Wilkinson, J.: Brief communication: New perspectives on the skill of modelled sea ice trends in light of recent Antarctic sea ice loss, The Cryosphere, 18, 5641–5652, https://doi.org/10.5194/tc-18-5641-2024, 2024.

Holt, L. A., Lott, F., Garcia, R. R., Kiladis, G. N., Cheng, Y., Anstey, J. A., Braesicke, P., Bushell, A. C., Butchart, N., Cagnazzo, C., Chen, C., Chun, H., Kawatani, Y., Kerzenmacher, T., Kim, Y., McLandress, C., Naoe, H., Osprey, S., Richter, J. H., Scaife, A. A., Scinocca, J., Serva, F., Versick, S., Watanabe, S., Yoshida, K., and Yukimoto, S.: An evaluation of tropical waves and wave forcing of the QBO in the QBOi models, Quarterly Journal of the Royal Meteorological Society, 148, 1541–1567, https://doi.org/10.1002/qj.3827, 2022.

Holton, J. R. and Tan, H.-C.: The Influence of the Equatorial Quasi-Biennial Oscillation on the Global Circulation at 50 mb, Journal of the Atmospheric Sciences, 37, 2200–2208, https://doi.org/10.1175/1520-0469(1980)037<2200:TIOTEQ>2.0.CO;2, 1980.

Hosseinzadehtalaei, P., Van Schaeybroeck, B., Termonia, P., and Tabari, H.: Identical hierarchy of physical drought types for climate change signals and uncertainty, Weather and Climate Extremes, 41, 100 573, https://doi.org/10.1016/j.wace.2023.100573, 2023.

Hou, M. and Tang, Y.: Recent progress in simulating two types of ENSO – from CMIP5 to CMIP6, Frontiers in Marine Science, 9, https://doi.org/10.3389/fmars.2022.986780, 2022.

Hou, Y., Guo, H., Yang, Y., and Liu, W.: Global evaluation of runoff simulation from climate, hydrological and land surface models, Water Resources Research, 59, e2021WR031 817, https://doi.org/https://doi.org/10.1029/2021WR031817, tex.eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2021WR031817, 2023.

Hu, D., Duan, A., Tang, Y., and Yu, W.: Delayed onset of the tropical Asian summer monsoon in CMIP6 can be linked to the cold bias over the Tibetan Plateau, Environmental Research Letters, 18, https://doi.org/10.1088/1748-9326/acff79, 2023.

Hu, Q., Li, T., Deng, X., Wu, T., Zhai, P., Huang, D., Fan, X., Zhu, Y., Lin, Y., Xiao, X., Chen, X., Zhao, X., Wang, L., and Qin, Z.: Intercomparison of global terrestrial carbon fluxes estimated by MODIS and Earth system models, Science of The Total Environment, 810, 152 231, https://doi.org/10.1016/j.scitotenv.2021.152231, 2022.

Huang, B., Stone, P. H., Sokolov, A. P., and Kamenkovich, I. V.: Ocean heat uptake in transient climate change: Mechanisms and uncertainty due to subgrid-scale eddy mixing, Journal of climate, 16, 3344–3356, 2003.

Hulme, M., Briffal, K. R., Jones, P. D., and Senior, C. A.: Validation of GCM control simulations using indices of daily airflow types over the British Isles, Climate Dynamics, 9, 95–105, https://doi.org/10.1007/BF00210012, 1993.

Huth, R., Beck, C., Philipp, A., Demuzere, M., Ustrnul, Z., Cahynová, M., Kyselý, J., and Tveito, O. E.: Classifications of Atmospheric Circulation Patterns: Recent Advances and Applications, Annals of the New York Academy of Sciences, 1146, 105–152, https://doi.org/10.1196/annals.1446.019, 2008.

Huusko, L. L., Bender, F. A.-M., Ekman, A. M. L., and Storelvmo, T.: Climate sensitivity indices and their relation with projected temperature change in CMIP6 models, Environmental Research Letters, 16, 064 095, https://doi.org/10.1088/1748-9326/ac0748, 2021.

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

IPCC: Climate Change 2021: The Physical Science Basis. Working Group I Contribution to the Sixth Assessment Report, https://www.ipcc.ch/report/ar6/wg1/, 2021.

Irving, D., Hobbs, W., Church, J., and Zika, J.: A Mass and Energy Conservation Analysis of Drift in the CMIP6 Ensemble, Journal of Climate, 34, 3157–3170, https://doi.org/10.1175/JCLI-D-20-0281.1, 2021.

Jackson, L. C., Biastoch, A., Buckley, M. W., Desbruyères, D. G., Frajka-Williams, E., Moat, B., and Robson, J.: The evolution of the North Atlantic Meridional Overturning Circulation since 1980, Nature Reviews Earth & Environment, 3, 241–254, https://doi.org/10.1038/s43017-022-00263-2, 2022.

James, P. M.: An objective classification method for Hess and Brezowsky Grosswetterlagen over Europe, Theoretical and Applied Climatology, 88, 17–42, https://doi.org/10.1007/s00704-006-0239-3, 2007.

Jensen, L., Eicker, A., Stacke, T., and Dobslaw, H.: Predictive Skill Assessment for Land Water Storage in CMIP5 Decadal Hindcasts by a Global Reconstruction of GRACE Satellite Data, Journal of Climate, 33, 9497–9509, https://doi.org/10.1175/JCLI-D-20-0042.1, 2020.

Jiang, J. H., Su, H., Wu, L., Zhai, C., and Schiro, K. A.: Improvements in Cloud and Water Vapor Simulations Over the Tropical Oceans in CMIP6 Compared to CMIP5, Earth and Space Science, 8, e2020EA001 520, https://doi.org/10.1029/2020EA001520, 2021.

Jiang, W., Gong, H., Huang, P., Wang, L., Huang, G., and Hu, L.: Biases and improvements of the ENSO-East Asian winter monsoon teleconnection in CMIP5 and CMIP6 models, Climate Dynamics, 59, 2467–2480, https://doi.org/10.1007/s00382-022-06220-5, 2022.

Jones, C. D. and Friedlingstein, P.: Quantifying process-level uncertainty contributions to TCRE and carbon budgets for meeting Paris Agreement climate targets, Environmental Research Letters, 15, 074 019, https://doi.org/10.1088/1748-9326/ab858a, 2020.

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, Geoscientific Model Development, 12, 4375–4385, https://doi.org/10.5194/gmd-12-4375-2019, 2019.

Jones, P. D., Hulme, M., and Briffa, K. R.: A comparison of Lamb circulation types with an objective classification scheme, International Journal of Climatology, 13, 655–663, https://doi.org/10.1002/joc.3370130606, 1993.

Jung, T., Balsamo, G., Bechtold, P., Beljaars, A. C. M., Köhler, M., Miller, M. J., Morcrette, J., Orr, A., Rodwell, M. J., and Tompkins, A. M.: The ECMWF model climate: recent progress through improved physical parametrizations, Quarterly Journal of the Royal Meteorological Society, 136, 1145–1160, https://doi.org/10.1002/qj.634, 2010.

Jung, T., Miller, M. J., Palmer, T. N., Towers, P., Wedi, N., Achuthavarier, D., Adams, J. M., Altshuler, E. L., Cash, B. A., Kinter, J. L., Marx, L., Stan, C., and Hodges, K. I.: High-Resolution Global Climate Simulations with the ECMWF Model in Project Athena: Experimental Design, Model Climate, and Seasonal Forecast Skill, Journal of Climate, 25, 3155–3172, https://doi.org/10.1175/JCLI-D-11-00265.1, 2012.

Kajtar, J. B., Santoso, A., McGregor, S., England, M. H., and Baillie, Z.: Model under-representation of decadal Pacific trade wind trends and its link to tropical Atlantic bias, Climate Dynamics, 50, 1471–1484, https://doi.org/10.1007/s00382-017-3699-5, 2018.

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Kaufman, Z., Wilson, E., Purich, A., Beadling, R., and Li, Y.: The Impact of Underestimated Southern Ocean Freshening on Simulated Historical Sea Surface Temperature Trends, Geophysical Research Letters, 52, e2024GL112 639, https://doi.org/10.1029/2024GL112639, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2024GL112639, 2025.

Kautz, L.-A., Martius, O., Pfahl, S., Pinto, J. G., Ramos, A. M., Sousa, P. M., and Woollings, T.: Atmospheric blocking and weather extremes over the Euro-Atlantic sector – a review, Weather and Climate Dynamics, 3, 305–336, https://doi.org/10.5194/wcd-3-305-2022, 2022.

Keen, A., Blockley, E., Bailey, D. A., Boldingh Debernard, J., Bushuk, M., Delhaye, S., Docquier, D., Feltham, D., Massonnet, F., O'Farrell, S., Ponsoni, L., Rodriguez, J. M., Schroeder, D., Swart, N., Toyoda, T., Tsujino, H., Vancoppenolle, M., and Wyser, K.: An inter-comparison of the mass budget of the Arctic sea ice in CMIP6 models, The Cryosphere, 15, 951–982, https://doi.org/10.5194/tc-15-951-2021, 2021.

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, Geoscientific Model Development, 11, 1133–1160, https://doi.org/10.5194/gmd-11-1133-2018, 2018.

Khosravi, N., Wang, Q., Koldunov, N., Hinrichs, C., Semmler, T., Danilov, S., and Jung, T.: The Arctic Ocean in CMIP6 Models: Biases and Projected Changes in Temperature and Salinity, Earth's Future, 10, e2021EF002 282, https://doi.org/10.1029/2021EF002282, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2021EF002282, 2022.

Kim, H., Lehner, F., Dagon, K., Lawrence, D. M., Swenson, S., and Wood, A. W.: Constraining climate model projections with observations amplifies future runoff declines, Communications Earth & Environment, 7, 182, https://doi.org/https://doi.org/10.1038/s43247-026-03213-8, 2026.

King, A. D., Sniderman, J. M. K., Dittus, A. J., Brown, J. R., Hawkins, E., and Ziehn, T.: Studying climate stabilization at Paris Agreement levels, Nature Climate Change, 11, 1010–1013, https://doi.org/10.1038/s41558-021-01225-0, 2021.

King, A. D., Abram, N. J., Alastrué de Asenjo, E., and Ziehn, T.: ESD Ideas: Extended net zero simulations are critical for informed decision making, Earth System Dynamics, 16, 1605–1609, https://doi.org/10.5194/esd-16-1605-2025, 2025.

Klein, S. A., Zhang, Y., Zelinka, M. D., Pincus, R., Boyle, J., and Gleckler, P. J.: Are climate model simulations of clouds improving? An evaluation using the ISCCP simulator, Journal of Geophysical Research: Atmospheres, 118, 1329–1342, 2013.

Kloster, S. and Lasslop, G.: Historical and future fire occurrence (1850 to 2100) simulated in CMIP5 Earth System Models, Global and Planetary Change, 150, 58–69, https://doi.org/10.1016/j.gloplacha.2016.12.017, 2017.

Knutti, R., Rugenstein, M. A. A., and Hegerl, G. C.: Beyond equilibrium climate sensitivity, Nature Geoscience, 10, 727–736, 2017.

Kociuba, G. and Power, S. B.: Inability of CMIP5 Models to Simulate Recent Strengthening of the Walker Circulation: Implications for Projections, Journal of Climate, 28, 20–35, https://doi.org/10.1175/JCLI-D-13-00752.1, 2015.

Kornhuber, K., Coumou, D., Vogel, E., Lesk, C., Donges, J. F., Lehmann, J., and Horton, R. M.: Amplified Rossby waves enhance risk of concurrent heatwaves in major breadbasket regions, Nature Climate Change, 10, 48–53, https://doi.org/10.1038/s41558-019-0637-z, 2020.

Koul, V., Tesdal, J.-E., Bersch, M., Hátún, H., Brune, S., Borchert, L., Haak, H., Schrum, C., and Baehr, J.: Unraveling the choice of the north Atlantic subpolar gyre index, Scientific Reports, 10, 1005, https://doi.org/10.1038/s41598-020-57790-5, 2020.

Koutsias, N., Xanthopoulos, G., Founda, D., Xystrakis, F., Nioti, F., Pleniou, M., Mallinis, G., and Arianoutsou, M.: On the relationships between forest fires and weather conditions in Greece from long-term national observations (1894–2010), International Journal of Wildland Fire, 22, 493–507, https://doi.org/10.1071/WF12003, 2013.

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Koven, C. D., Hugelius, G., Lawrence, D. M., and Wieder, W. R.: Higher climatological temperature sensitivity of soil carbon in cold than warm climates, Nature Climate Change, 7, 817–822, https://doi.org/10.1038/nclimate3421, 2017.

Kramer, R., Smith, C., and Andrews, T.: The radiative forcing model intercomparison project (RFMIP2. 0) for CMIP7, EGUsphere, 2025, 1–30, 2025.

Kroll, C. A., Jnglin Wills, R. C., Kornblueh, L., Niemeier, U., and Schneidereit, A.: Parameterization adaption needed to unlock the benefits of increased resolution for the ITCZ in ICON, Atmospheric Chemistry and Physics, 25, 16 915–16 943, https://doi.org/10.5194/acp-25-16915-2025, 2025.

Kusahara, K. and Tatebe, H.: Causes of the Abrupt and Sustained 2016–2023 Antarctic Sea-Ice Decline: A Sea Ice–Ocean Model Perspective, Geophysical Research Letters, 52, e2025GL115 256, https://doi.org/https://doi.org/10.1029/2025GL115256, 2025.

Lauer, A., Jones, C., Eyring, V., Evaldsson, M., Hagemann, S., Mäkelä, J., Martin, G., Roehrig, R., and Wang, S.: Process-level improvements in CMIP5 models and their impact on tropical variability, the Southern Ocean, and monsoons, Earth System Dynamics, 9, 33–67, 2018.

Lauer, A., Bock, L., Hassler, B., Schröder, M., and Stengel, M.: Cloud Climatologies from Global Climate Models—A Comparison of CMIP5 and CMIP6 Models with Satellite Data, Journal of Climate, 36, 281–311, https://doi.org/10.1175/JCLI-D-22-0181.1, 2023.

Le, P. V. V., Guilloteau, C., Mamalakis, A., and Foufoula-Georgiou, E.: Underestimated MJO Variability in CMIP6 Models, Geophysical Research Letters, 48, e2020GL092 244, https://doi.org/10.1029/2020GL092244, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020GL092244, 2021.

Lee, J., Sperber, K. R., Gleckler, P. J., Bonfils, C. J. W., and Taylor, K. E.: Quantifying the agreement between observed and simulated extratropical modes of interannual variability, Climate Dynamics, 52, 4057–4089, https://doi.org/10.1007/s00382-018-4355-4, 2019.

Lee, J., Planton, Y. Y., Gleckler, P. J., Sperber, K. R., Guilyardi, E., Wittenberg, A. T., McPhaden, M. J., and Pallotta, G.: Robust Evaluation of ENSO in Climate Models: How Many Ensemble Members Are Needed?, Geophysical Research Letters, 48, e2021GL095 041, https://doi.org/10.1029/2021GL095041, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2021GL095041, 2021a.

Lee, J., Sperber, K. R., Gleckler, P. J., Taylor, K. E., and Bonfils, C. J. W.: Benchmarking performance changes in the simulation of extratropical modes of variability across CMIP generations, Journal of Climate, pp. 1–70, https://doi.org/10.1175/JCLI-D-20-0832.1, 2021b.

Lehner, F., Wood, A. W., Vano, J. A., Lawrence, D. M., Clark, M. P., and Mankin, J. S.: The potential to reduce uncertainty in regional runoff projections from climate models, Nature Climate Change, 9, 926–933, https://doi.org/https://doi.org/10.1038/s41558-019-0639-x, 2019.

Li, F., Levis, S., and Ward, D. S.: Quantifying the role of fire in the Earth system – Part 1: Improved global fire modeling in the Community Earth System Model (CESM1), Biogeosciences, 10, 2293–2314, https://doi.org/10.5194/bg-10-2293-2013, 2013.

Li, F., Song, X., Harrison, S. P., Marlon, J. R., Lin, Z., Leung, L. R., Schwinger, J., Marécal, V., Wang, S., Ward, D. S., Dong, X., Lee, H., Nieradzik, L., Rabin, S. S., and Séférian, R.: Evaluation of global fire simulations in CMIP6 Earth system models, Geoscientific Model Development, 17, 8751–8771, https://doi.org/10.5194/gmd-17-8751-2024, 2024a.

Li, J., Miao, C., Wei, W., Zhang, G., Hua, L., Chen, Y., and Wang, X.: Evaluation of CMIP6 Global Climate Models for Simulating Land Surface Energy and Water Fluxes During 1979–2014, Journal of Advances in Modeling Earth Systems, 13, https://doi.org/10.1029/2021MS002515, 2021.

Li, X., Wolf, M., Shen, X., Steinke, I., Lai, Z., Niu, S., China, S., Shrivastava, M., Zhang, Z., Gold, A., and others: Quantifying and modeling the impact of phase state on the ice nucleation abilities of 2-methyltetrols as a key component of secondary organic aerosol derived from isoprene epoxydiols, Environmental Science & Technology, 58, 22 678–22 690, 2024b.

Li, X., Fan, S., Guo, H., and Ginoux, P.: Resolving the roles of soot and dust in cirrus cloud ice formation at regional and global scales: insights from parcel and climate modeling, Atmospheric Chemistry and Physics, 26, 95–115, 2026.

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Li, Z., Zou, L., Zhao, A., Zhang, H., Yue, F., Luo, Z., Bian, R., and Xu, R.: Evaluation of annual trends in carbon cycle variables simulated by CMIP6 Earth system models in China, Geoscientific Model Development, 18, 8703–8722, https://doi.org/10.5194/gmd-18-8703-2025, 2025.

Liao, H., Wang, C., and Song, Z.: ENSO phase-locking biases from the CMIP5 to CMIP6 models and a possible explanation, Deep Sea Research Part II: Topical Studies in Oceanography, 189-190, 104 943, https://doi.org/10.1016/j.dsr2.2021.104943, 2021.

Lieber, R., King, A., Brown, J., Ashcroft, L., Freund, M., and McMichael, C.: ENSO Teleconnections More Uncertain in Regions of Lower Socioeconomic Development, Geophysical Research Letters, 49, e2022GL100 553, https://doi.org/10.1029/2022GL100553, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2022GL100553, 2022.

Lieber, R., Brown, J., King, A., and Freund, M.: Historical and Future Asymmetry of ENSO Teleconnections with Extremes, Journal of Climate, 37, 5909–5924, https://doi.org/10.1175/JCLI-D-23-0619.1, 2024.

Lindzen, R. S. and Holton, J. R.: A Theory of the Quasi-Biennial Oscillation, Journal of the Atmospheric Sciences, 25, 1095–1107, https://doi.org/10.1175/1520-0469(1968)025<1095:ATOTQB>2.0.CO;2, 1968.

Littell, J. S., Peterson, D. L., Riley, K. L., Liu, Y., and Luce, C. H.: A review of the relationships between drought and forest fire in the United States, Global Change Biology, 22, 2353–2369, https://doi.org/10.1111/gcb.13275, 2016.

Liu, Y., Cheng, L., Pan, Y., Tan, Z., Abraham, J., Zhang, B., Zhu, J., and Song, J.: How Well Do CMIP6 and CMIP5 Models Simulate the Climatological Seasonal Variations in Ocean Salinity?, Advances in Atmospheric Sciences, 39, 1650–1672, https://doi.org/10.1007/s00376-022-1381-2, 2022.

Liu, Y., Cai, W., Lin, X., Li, Z., and Zhang, Y.: Nonlinear El Niño impacts on the global economy under climate change, Nature Communications, 14, 5887, https://doi.org/10.1038/s41467-023-41551-9, 2023.

Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J. G., Liang, L., Mitrescu, C., Rose, F. G., and Kato, S.: Clouds and the earth's radiant energy system (CERES) energy balanced and filled (EBAF) top-of-atmosphere (TOA) edition-4.0 data product, Journal of climate, 31, 895–918, 2018.

Lozier, M. S., Li, F., Bacon, S., Bahr, F., Bower, A. S., Cunningham, S. A., Jong, M. F. d., Steur, L. d., deYoung, B., Fischer, J., Gary, S. F., Greenan, B. J. W., Holliday, N. P., Houk, A., Houpert, L., Inall, M. E., Johns, W. E., Johnson, H. L., Johnson, C., Karstensen, J., Koman, G., Bras, I. A. L., Lin, X., Mackay, N., Marshall, D. P., Mercier, H., Oltmanns, M., Pickart, R. S., Ramsey, A. L., Rayner, D., Straneo, F., Thierry, V., Torres, D. J., Williams, R. G., Wilson, C., Yang, J., Yashayaev, I., and Zhao, J.: A sea change in our view of overturning in the subpolar North Atlantic, Science, 363, 516–521, https://doi.org/10.1126/science.aau6592, 2019.

Luo, F., Selten, F., Wehrli, K., Kornhuber, K., Le Sager, P., May, W., Reerink, T., Seneviratne, S. I., Shiogama, H., Tokuda, D., Kim, H., and Coumou, D.: Summertime Rossby waves in climate models: substantial biases in surface imprint associated with small biases in upper-level circulation, Weather and Climate Dynamics, 3, 905–935, https://doi.org/10.5194/wcd-3-905-2022, 2022.

Ma, X., Zhao, S., Zhang, H., and Wang, W.: The double-ITCZ problem in CMIP6 and the influences of deep convection and model resolution, International Journal of Climatology, 43, 2369–2390, https://doi.org/10.1002/joc.7980, 2023.

MacDougall, A. H.: The Transient Response to Cumulative CO2 Emissions: a Review, Current Climate Change Reports, 2, 39–47, https://doi.org/10.1007/s40641-015-0030-6, 2016.

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., Mengis, N., 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

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

there warming in the pipeline? A multi-model analysis of the Zero Emissions Commitment from CO_2, Biogeosciences, 17, 2987–3016, https://doi.org/10.5194/bg-17-2987-2020, 2020.

Maddison, J. W., Catto, J. L., Hanna, E., Luu, L. N., and Screen, J. A.: Missing Increase in Summer Greenland Blocking in Climate Models, Geophysical Research Letters, 51, e2024GL108 505, https://doi.org/10.1029/2024GL108505, 2024.

Maher, N., Gupta, A. S., and England, M. H.: Drivers of decadal hiatus periods in the 20th and 21st centuries, Geophysical Research Letters, 41, 5978–5986, https://doi.org/10.1002/2014GL060527, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2014GL060527, 2014.

Maher, N., Phillips, A. S., Deser, C., Wills, R. C. J., Lehner, F., Fasullo, J., Caron, J. M., Brunner, L., Beyerle, U., and Jeffree, J.: The updated Multi-Model Large Ensemble Archive and the Climate Variability Diagnostics Package: new tools for the study of climate variability and change, Geoscientific Model Development, 18, 6341–6365, https://doi.org/10.5194/gmd-18-6341-2025, 2025.

Mansfield, L. A. and Sheshadri, A.: Uncertainty Quantification of a Machine Learning Subgrid-Scale Parameterization for Atmospheric Gravity Waves, Journal of Advances in Modeling Earth Systems, 16, e2024MS004 292, https://doi.org/10.1029/2024MS004292, 2024.

Mantua, N. J., Hare, S. R., Zhang, Y., Wallace, J. M., and Francis, R. C.: A Pacific Interdecadal Climate Oscillation with Impacts on Salmon Production, Bulletin of the American Meteorological Society, 78, 1069–1079, https://doi.org/10.1175/1520-0477(1997)078<1069:APICOW>2.0.CO;2, 1997.

Maraun, D., Shepherd, T. G., Widmann, M., Zappa, G., Walton, D., Gutiérrez, J. M., Hagemann, S., Richter, I., Soares, P. M. M., Hall, A., and Mearns, L. O.: Towards process-informed bias correction of climate change simulations, Nature Climate Change, 7, 764–773, https://doi.org/10.1038/nclimate3418, 2017.

Martin, T. and Biastoch, A.: On the ocean's response to enhanced Greenland runoff in model experiments: relevance of mesoscale dynamics and atmospheric coupling, Ocean Science, 19, 141–167, https://doi.org/10.5194/os-19-141-2023, 2023.

Matthes, H., Damseaux, A., Westermann, S., Beer, C., Boone, A., Burke, E., Decharme, B., Genet, H., Jafarov, E., Langer, M., Parmentier, F.-J., Porada, P., Gagne-Landmann, A., Huntzinger, D., Rogers, B. M., Schädel, C., Stacke, T., Wells, J., and Wieder, W. R.: Advances in Permafrost Representation: Biophysical Processes in Earth System Models and the Role of Offline Models, Permafrost and Periglacial Processes, 36, 302–318, https://doi.org/10.1002/ppp.2269, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ppp.2269, 2025.

Matveeva, T. A. and Semenov, V. A.: Regional Features of the Arctic Sea Ice Area Changes in 2000–2019 versus 1979–1999 Periods, Atmosphere, 13, 1434, https://doi.org/10.3390/atmos13091434, 2022.

McCarthy, G. D. and Caesar, L.: Can we trust projections of AMOC weakening based on climate models that cannot reproduce the past?, Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences, 381, 20220 193, https://doi.org/10.1098/rsta.2022.0193, 2023.

McGregor, S., Timmermann, A., Stuecker, M. F., England, M. H., Merrifield, M., Jin, F.-F., and Chikamoto, Y.: Recent Walker circulation strengthening and Pacific cooling amplified by Atlantic warming, Nature Climate Change, 4, 888–892, https://doi.org/10.1038/nclimate2330, 2014.

McGregor, S., Stuecker, M. F., Kajtar, J. B., England, M. H., and Collins, M.: Model tropical Atlantic biases underpin diminished Pacific decadal variability, Nature Climate Change, 8, 493–498, https://doi.org/10.1038/s41558-018-0163-4, 2018.

McKenna, S., Santoso, A., Gupta, A. S., Taschetto, A. S., and Cai, W.: Indian Ocean Dipole in CMIP5 and CMIP6: characteristics, biases, and links to ENSO, Scientific Reports, 10, 11 500, https://doi.org/10.1038/s41598-020-68268-9, 2020.

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

McKenna, S., Santoso, A., Sen Gupta, A., and Taschetto, A. S.: Understanding Biases in Indian Ocean Seasonal SST in CMIP6 Models, Journal of Geophysical Research: Oceans, 129, e2023JC020 330, https://doi.org/10.1029/2023JC020330, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2023JC020330, 2024.

McPhaden, M. J., Zebiak, S. E., and Glantz, M. H.: ENSO as an Integrating Concept in Earth Science, Science, 314, 1740–1745, https://doi.org/10.1126/science.1132588, 2006.

Meehl, G. A.: The South Asian Monsoon and the Tropospheric Biennial Oscillation, Journal of Climate, 10, 1921–1943, https://doi.org/10.1175/1520-0442(1997)010<1921:TSAMAT>2.0.CO;2, 1997.

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, https://doi.org/10.1126/sciadv.aba1981, tex.eprint: https://www.science.org/doi/pdf/10.1126/sciadv.aba1981, 2020.

Meier, W. N., Stroeve, J., and Fetterer, F.: Whither Arctic sea ice? A clear signal of decline regionally, seasonally and extending beyond the satellite record, Annals of Glaciology, 46, 428–434, https://doi.org/10.3189/172756407782871170, 2007.

Meier, W. N., Stewart, J. S., Windnagel, A., and Fetterer, F. M.: Comparison of hemispheric and regional sea ice extent and area trends from NOAA and NASA passive microwave-derived climate records, Remote Sensing, 14, 619, 2022.

Menary, M. B., Robson, J., Allan, R. P., Booth, B. B. B., Cassou, C., Gastineau, G., Gregory, J., Hodson, D., Jones, C., Mignot, J., Ringer, M., Sutton, R., Wilcox, L., and Zhang, R.: Aerosol-Forced AMOC Changes in CMIP6 Historical Simulations, Geophysical Research Letters, 47, e2020GL088 166, https://doi.org/10.1029/2020GL088166, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020GL088166, 2020.

Miao, C., Wu, Y., Fan, X., and Su, J.: Projections of global land runoff changes and their uncertainty characteristics during the 21st century, Earth's Future, 11, e2022EF003 286, https://doi.org/https://doi.org/10.1029/2022EF003286, tex.eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2022EF003286, 2023.

Minière, A., von Schuckmann, K., Sallée, J.-B., and Vogt, L.: Robust acceleration of Earth system heating observed over the past six decades, Scientific Reports, 13, 22 975, https://doi.org/10.1038/s41598-023-49353-1, 2023.

Mishra, A. N., Maraun, D., Schiemann, R., Hodges, K., Zappa, G., and Ossó, A.: Long-lasting intense cut-off lows to become more frequent in the Northern Hemisphere, Communications Earth & Environment, 6, 115, https://doi.org/10.1038/s43247-025-02078-7, 2025.

Moat, B. I.: Atlantic meridional overturning circulation observed by the RAPID-MOCHA-WBTS array at 26N from 2004 to 2024 (v2024.1), https://doi.org/10.5285/3f24651e-2d44-dee3-e063-7086abc0395e, archive Location: North Atlantic Ocean,Straits of Florida,Atlantic Ocean,Northeast Atlantic Ocean (40W),Northwest Atlantic Ocean (40W), 2025.

Mohan, S. and Ruchith, R.: On the simulations of latent heat flux over the Indian Ocean in CMIP6 models, Climate Dynamics, 61, 5573–5595, https://doi.org/10.1007/s00382-023-06871-y, 2023.

Moreno-Chamarro, E., Caron, L.-P., Loosveldt Tomas, S., Vegas-Regidor, J., Gutjahr, O., Moine, M.-P., Putrasahan, D., Roberts, C., Roberts, M., Senan, R., Terray, L., Tourigny, E., and Vidale, P.: Impact of increased resolution on long-standing biases in HighResMIP-PRIMAVERA climate models, Geoscientific Model Development, 15, 269–289, https://doi.org/10.5194/gmd-15-269-2022, 2022.

Mueller, B. and Seneviratne, S. I.: Systematic land climate and evapotranspiration biases in CMIP5 simulations, Geophysical Research Letters, 41, 128–134, https://doi.org/10.1002/2013GL058055, 2014.

Mulcahy, J., Jones, C., Rumbold, S., Kuhlbrodt, T., Dittus, A., Blockley, E., Yool, A., Walton, J., Hardacre, C., Andrews, T., Bodas-Salcedo, A., Stringer, M., De Mora, L., Harris, P., Hill, R., Kelley, D., Robertson, E., and Tang, Y.: UKESM1.1: development and evaluation of an

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

updated configuration of the UK Earth System Model, Geoscientific Model Development, 16, 1569–1600, https://doi.org/10.5194/gmd-16-1569-2023, 2023.

Nakamura, N. and Huang, C. S. Y.: Atmospheric blocking as a traffic jam in the jet stream, Science, 361, 42–47, https://doi.org/10.1126/science.aat0721, 2018.

Nam, C., Bony, S., Dufresne, J.-L., and Chepfer, H.: The 'too few, too bright'tropical low-cloud problem in CMIP5 models, Geophysical Research Letters, 39, 2012.

Nie, Y., Lin, X., Yang, Q., Liu, J., Chen, D., and Uotila, P.: Differences between the CMIP5 and CMIP6 Antarctic sea ice concentration budgets, Geophysical Research Letters, 50, e2023GL105 265, 2023.

Nieto, R., Gimeno, L., De La Torre, L., Ribera, P., Gallego, D., García-Herrera, R., García, J. A., Nuñez, M., Redaño, A., and Lorente, J.: Climatological Features of Cutoff Low Systems in the Northern Hemisphere, Journal of Climate, 18, 3085–3103, https://doi.org/10.1175/JCLI3386.1, 2005.

Notz, D. and Community, S.: Arctic Sea Ice in CMIP6, Geophysical Research Letters, 47, https://doi.org/10.1029/2019gl086749, 2020.

Osborn, T., Conway, D., Hulme, M., Gregory, J., and Jones, P.: Air flow influences on local climate: observed and simulated mean relationships for the United Kingdom, Climate Research, 13, 173–191, https://doi.org/10.3354/cr013173, 1999.

O'Sullivan, M., Smith, W. K., Sitch, S., Friedlingstein, P., Arora, V. K., Haverd, V., Jain, A. K., Kato, E., Kautz, M., Lombardozzi, D., Nabel, J. E. M. S., Tian, H., Vuichard, N., Wiltshire, A., Zhu, D., and Buermann, W.: Climate-Driven Variability and Trends in Plant Productivity Over Recent Decades Based on Three Global Products, Global Biogeochemical Cycles, 34, e2020GB006 613, https://doi.org/10.1029/2020GB006613, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020GB006613, 2020.

Otero, N., Sillmann, J., and Butler, T.: Assessment of an extended version of the Jenkinson–Collison classification on CMIP5 models over Europe, Climate Dynamics, 50, 1559–1579, https://doi.org/10.1007/s00382-017-3705-y, 2018.

Oueslati, B. and Bellon, G.: The double ITCZ bias in CMIP5 models: Interaction between SST, large-scale circulation and precipitation, Climate dynamics, 44, 585–607, 2015.

O'Gorman, P. A. and Schneider, T.: The Hydrological Cycle over a Wide Range of Climates Simulated with an Idealized GCM, Journal of Climate, 21, 3815–3832, https://doi.org/10.1175/2007JCLI2065.1, 2008.

O'Reilly, C. H., Minobe, S., and Kuwano-Yoshida, A.: The influence of the Gulf Stream on wintertime European blocking, Climate Dynamics, 47, 1545–1567, https://doi.org/10.1007/s00382-015-2919-0, 2016.

Palazzo Corner, S., Siegert, M., Ceppi, P., Fox-Kemper, B., Frölicher, T. L., Gallego-Sala, A., Haigh, J., Hegerl, G. C., Jones, C. D., Knutti, R., Koven, C. D., MacDougall, A. H., Meinshausen, M., Nicholls, Z., Sallée, J. B., Sanderson, B. M., Séférian, R., Turetsky, M., Williams, R. G., Zaehle, S., and Rogelj, J.: The Zero Emissions Commitment and climate stabilization, Frontiers in Science, 1, https://doi.org/10.3389/fsci.2023.1170744, 2023.

Pang, S., Wang, X., and Vialard, J.: How Well Do CMIP6 Models Simulate Salinity Barrier Layers in the North Indian Ocean?, Journal of Climate, 37, 289–308, https://doi.org/10.1175/JCLI-D-23-0366.1, 2024.

Papadimitriou, L. V., Koutroulis, A. G., Grillakis, M. G., and Tsanis, I. K.: The effect of GCM biases on global runoff simulations of a land surface model, Hydrology and Earth System Sciences, 21, 4379–4401, https://doi.org/10.5194/hess-21-4379-2017, 2017.

Park, H. and Jeong, S.: Leaf area index in Earth system models: how the key variable of vegetation seasonality works in climate projections, Environmental Research Letters, 16, 034 027, https://doi.org/10.1088/1748-9326/abe2cf, 2021.

Park, J.-Y., Schloesser, F., Timmermann, A., Choudhury, D., Lee, J.-Y., and Nellikkattil, A. B.: Future sea-level projections with a coupled atmosphere-ocean-ice-sheet model, Nature Communications, 14, 636, https://doi.org/10.1038/s41467-023-36051-9, 2023.

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Parkinson, C. L.: A 40-y record reveals gradual Antarctic sea ice increases followed by decreases at rates far exceeding the rates seen in the Arctic, Proceedings of the National Academy of Sciences, 116, 14 414–14 423, https://doi.org/10.1073/pnas.1906556116, 2019.

Perez, J., Menendez, M., Mendez, F. J., and Losada, I. J.: Evaluating the performance of CMIP3 and CMIP5 global climate models over the north-east Atlantic region, Climate Dynamics, 43, 2663–2680, https://doi.org/10.1007/s00382-014-2078-8, 2014.

Perkins-Kirkpatrick, S., Palmer, L., King, A., and Ziehn, T.: Heatwaves in a net zero World, Environmental Research: Climate, 4, 045 015, https://doi.org/10.1088/2752-5295/ae0ea4, 2025.

Petoukhov, V., Rahmstorf, S., Petri, S., and Schellnhuber, H. J.: Quasiresonant amplification of planetary waves and recent Northern Hemisphere weather extremes, Proceedings of the National Academy of Sciences, 110, 5336–5341, https://doi.org/10.1073/pnas.1222000110, 2013.

Pinheiro, H., Ambrizzi, T., Hodges, K., Gan, M., Andrade, K., and Garcia, J.: Are Cut-off Lows simulated better in CMIP6 compared to CMIP5?, Climate Dynamics, 59, 2117–2136, https://doi.org/10.1007/s00382-022-06200-9, 2022.

Planton, Y. Y., Guilyardi, E., Wittenberg, A. T., Lee, J., Gleckler, P. J., Bayr, T., McGregor, S., McPhaden, M. J., Power, S., Roehrig, R., Vialard, J., and Voldoire, A.: Evaluating Climate Models with the CLIVAR 2020 ENSO Metrics Package, Bulletin of the American Meteorological Society, 102, E193–E217, https://doi.org/10.1175/BAMS-D-19-0337.1, 2021a.

Planton, Y. Y., Guilyardi, E., Wittenberg, A. T., Lee, J., Gleckler, P. J., Bayr, T., McGregor, S., McPhaden, M. J., Power, S., Roehrig, R., Vialard, J., and Voldoire, A.: A New Way of Evaluating ENSO in Climate Models: The CLIVAR ENSO Metrics Package, Bulletin of the American Meteorological Society, 102, 1073–1080, https://doi.org/10.1175/BAMS-D-19-0337.A, 2021b.

Portmann, V.: GMD - ClimLoco1.0: CLimate variable confidence Interval of Multivariate Linear Observational COnstraint, https://gmd.copernicus.org/articles/18/9015/2025/, 2025.

Power, S., Casey, T., Folland, C., Colman, A., and Mehta, V.: Inter-decadal modulation of the impact of ENSO on Australia, Climate Dynamics, 15, 319–324, https://doi.org/10.1007/s003820050284, 1999.

Power, S., Lengaigne, M., Capotondi, A., Khodri, M., Vialard, J., Jebri, B., Guilyardi, E., McGregor, S., Kug, J.-S., Newman, M., McPhaden, M. J., Meehl, G., Smith, D., Cole, J., Emile-Geay, J., Vimont, D., Wittenberg, A. T., Collins, M., Kim, G.-I., Cai, W., Okumura, Y., Chung, C., Cobb, K. M., Delage, F., Planton, Y. Y., Levine, A., Zhu, F., Sprintall, J., Di Lorenzo, E., Zhang, X., Luo, J.-J., Lin, X., Balmaseda, M., Wang, G., and Henley, B. J.: Decadal climate variability in the tropical Pacific: Characteristics, causes, predictability, and prospects, Science, 374, eaay9165, https://doi.org/10.1126/science.aay9165, 2021.

Priya, P. and Dommenget, D.: ENSO phase space dynamics in CMIP models, Climate Dynamics, 63, 270, https://doi.org/10.1007/s00382-025-07752-2, 2025.

Purich, A. and Doddridge, E. W.: Record low Antarctic sea ice coverage indicates a new sea ice state, Communications Earth & Environment, 4, 1–9, https://doi.org/10.1038/s43247-023-00961-9, 2023.

Ramanathan, V., Cess, R., Harrison, E., Minnis, P., Barkstrom, B., Ahmad, E., and Hartmann, D.: Cloud-radiative forcing and climate: Results from the earth radiation budget experiment, Science, 243, 57–63, 1989.

Reddy, P. J., Perkins-Kirkpatrick, S. E., and Sharples, J. J.: Interactive influence of ENSO and IOD on contiguous heatwaves in Australia, Environmental Research Letters, 17, 014 004, https://doi.org/10.1088/1748-9326/ac3e9a, 2021.

Reed, R. J., Campbell, W. J., Rasmussen, L. A., and Rogers, D. G.: Evidence of a downward-propagating, annual wind reversal in the equatorial stratosphere, Journal of Geophysical Research (1896-1977), 66, 813–818, https://doi.org/10.1029/JZ066i003p00813, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/JZ066i003p00813, 1961.

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Reintges, A., Robson, J. I., Sutton, R., and Yeager, S. G.: Subpolar North Atlantic Mean State Affects the Response of the Atlantic Meridional Overturning Circulation to the North Atlantic Oscillation in CMIP6 Models, Journal of Climate, 37, 5543–5559, https://doi.org/10.1175/JCLI-D-23-0470.1, 2024.

Richter, I.: Climate model biases in the eastern tropical oceans: causes, impacts and ways forward, WIREs Climate Change, 6, 345–358, https://doi.org/10.1002/wcc.338, _eprint: https://wires.onlinelibrary.wiley.com/doi/pdf/10.1002/wcc.338, 2015.

Richter, I. and Tokinaga, H.: An overview of the performance of CMIP6 models in the tropical Atlantic: mean state, variability, and remote impacts, Climate Dynamics, 55, 2579–2601, https://doi.org/10.1007/s00382-020-05409-w, 2020.

Richter, I., Chang, P., Chiu, P.-G., Danabasoglu, G., Doi, T., Dommenget, D., Gastineau, G., Gillett, Z. E., Hu, A., Kataoka, T., Keenlyside, N. S., Kucharski, F., Okumura, Y. M., Park, W., Stuecker, M. F., Taschetto, A. S., Wang, C., Yeager, S. G., and Yeh, S.-W.: The Tropical Basin Interaction Model Intercomparison Project (TBIMIP), Geoscientific Model Development, 18, 2587–2608, https://doi.org/10.5194/gmd-18-2587-2025, 2025.

Rivera, J. A. and Arnould, G.: Evaluation of the ability of CMIP6 models to simulate precipitation over Southwestern South America: Climatic features and long-term trends (1901–2014), Atmospheric Research, 241, 104 953, https://doi.org/10.1016/j.atmosres.2020.104953, 2020.

Roach, L. A., Dean, S. M., and Renwick, J. A.: Consistent biases in Antarctic sea ice concentration simulated by climate models, The Cryosphere, 12, 365–383, 2018.

Roach, L. A., Dörr, J., Holmes, C. R., Massonnet, F., Blockley, E. W., Notz, D., Rackow, T., Raphael, M. N., O'Farrell, S. P., Bailey, D. A., and Bitz, C. M.: Antarctic Sea Ice Area in CMIP6, Geophysical Research Letters, 47, e2019GL086 729, https://doi.org/10.1029/2019GL086729, 2020.

Roach, L. A., Eisenman, I., Wagner, T. J. W., Blanchard-Wrigglesworth, E., and Bitz, C. M.: Asymmetry in the seasonal cycle of Antarctic sea ice driven by insolation, Nature Geoscience, 15, 277–281, https://doi.org/10.1038/s41561-022-00913-6, 2022.

Roberts, M. J., Jackson, L. C., Roberts, C. D., Meccia, V., Docquier, D., Koenigk, T., Ortega, P., Moreno-Chamarro, E., Bellucci, A., Coward, A., Drijfhout, S., Exarchou, E., Gutjahr, O., Hewitt, H., Iovino, D., Lohmann, K., Putrasahan, D., Schiemann, R., Seddon, J., Terray, L., Xu, X., Zhang, Q., Chang, P., Yeager, S. G., Castruccio, F. S., Zhang, S., and Wu, L.: Sensitivity of the Atlantic Meridional Overturning Circulation to Model Resolution in CMIP6 HighResMIP Simulations and Implications for Future Changes, Journal of Advances in Modeling Earth Systems, 12, e2019MS002 014, https://doi.org/10.1029/2019MS002014, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019MS002014, 2020.

Robson, J., Sutton, R., Menary, M. B., and Lai, M. W. K.: Contrasting internally and externally generated Atlantic Multidecadal Variability and the role for AMOC in CMIP6 historical simulations, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 381, 20220 194, https://doi.org/10.1098/rsta.2022.0194, 2023.

Rustiana, S., Trismidianto, and Satyawardhana, H.: The Influence of ENSO and IOD During Mesoscale Convective Complex (MCC) to Rainfall in Indonesia, IOP Conference Series: Earth and Environmental Science, 303, 012 006, https://doi.org/10.1088/1755-1315/303/1/012006, 2019.

Samuels, M., Adam, O., and Gildor, H.: A Shallow Thermocline Bias in the Southern Tropical Pacific in CMIP5/6 Models Linked to Double-ITCZ Bias, Geophysical Research Letters, 48, e2021GL093 818, https://doi.org/10.1029/2021GL093818, 2021.

Sanderson, B. M., Booth, B. B. B., Dunne, J., Eyring, V., Fisher, R. A., Friedlingstein, P., Gidden, M. J., Hajima, T., Jones, C. D., Jones, C. G., King, A., Koven, C. D., Lawrence, D. M., Lowe, J., Mengis, N., Peters, G. P., Rogelj, J., Smith, C., Snyder, A. C., Simpson, I. R., Swann, A. L. S., Tebaldi, C., Ilyina, T., Schleussner, C.-F., Séférian, R., Samset, B. H., van Vuuren, D., and Zaehle, S.: The need for carbon-emissions-driven climate projections in CMIP7, Geoscientific Model Development, 17, 8141–8172, https://doi.org/10.5194/gmd-17-8141-2024, 2024.

Sanderson, B. M., Brovkin, V., Fisher, R. A., Hohn, D., Ilyina, T., Jones, C. D., Koenigk, T., Koven, C., Li, H., Lawrence, D. M., Lawrence, P., Liddicoat, S., MacDougall, A. H., Mengis, N., Nicholls, Z., O'Rourke, E., Romanou, A., Sandstad, M., Schwinger, J., Séférian, R., Sentman, L. T., Simpson, I. R., Smith, C., Steinert, N. J., Swann, A. L. S., Tjiputra, J., and Ziehn, T.: flat10MIP: an emissions-driven experiment to diagnose the climate response to positive, zero and negative CO_2 emissions, Geoscientific Model Development, 18, 5699–5724, https://doi.org/10.5194/gmd-18-5699-2025, 2025.

Sang, Y., Ren, H.-L., Shi, X., Xu, X., and Chen, H.: Improvement of Soil Moisture Simulation in Eurasia by the Beijing Climate Center Climate System Model from CMIP5 to CMIP6, Advances in Atmospheric Sciences, 38, 237–252, https://doi.org/10.1007/s00376-020-0167-7, 2021.

Santoso, A., Cai, W., Collins, M., McPhaden, M., Jin, F.-F., Guilyardi, E., Vecchi, G., Dommenget, D., and Wang, G.: ENSO EXTREMES AND DIVERSITY: Dynamics, Teleconnections, and Impacts, Bulletin of the American Meteorological Society, 96, 1969–1972, https://www.jstor.org/stable/26225047, 2015.

Santoso, A., Taschetto, A. S., McGregor, S., Roxy, M. K., Chung, C., Wu, B., and Delage, F. P.: Dynamics and impacts of tropical climate variability: Understanding trends and future projections, https://www.frontiersin.org/articles/10.3389/fclim.2023.1148145/full, pages: 1148145 Publication Title: Frontiers in Climate Volume: 5, 2023.

Scaife, A. A., Copsey, D., Gordon, C., Harris, C., Hinton, T., Keeley, S., O'Neill, A., Roberts, M., and Williams, K.: Improved Atlantic winter blocking in a climate model: ATLANTIC BLOCKING, Geophysical Research Letters, 38, n/a–n/a, https://doi.org/10.1029/2011GL049573, 2011.

Schiemann, R., Demory, M.-E., Shaffrey, L. C., Strachan, J., Vidale, P. L., Mizielinski, M. S., Roberts, M. J., Matsueda, M., Wehner, M. F., and Jung, T.: The Resolution Sensitivity of Northern Hemisphere Blocking in Four 25-km Atmospheric Global Circulation Models, Journal of Climate, 30, 337–358, https://doi.org/10.1175/JCLI-D-16-0100.1, 2017.

Schiller-Weiss, I., Martin, T., and Schwarzkopf, F. U.: Emerging Influence of Enhanced Greenland Melting on Boundary Currents and Deep Convection Regimes in the Labrador and Irminger Seas, Geophysical Research Letters, 51, e2024GL109 022, https://doi.org/10.1029/2024GL109022, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2024GL109022, 2024.

Schroeter, S.: Seasonal and Regional Antarctic Sea Ice Biases: A Closer Look at CMIP6, Atmosphere-Ocean, 63, 183–200, https://doi.org/10.1080/07055900.2025.2507880, 2025.

Segura, H., Pedruzo-Bagazgoitia, X., Weiss, P., Müller, S. K., Rackow, T., Lee, J., Dolores-Tesillos, E., Benedict, I., Aengenheyster, M., Aguridan, R., Arduini, G., Baker, A. J., Bao, J., Bastin, S., Baulenas, E., Becker, T., Beyer, S., Bockelmann, H., Brüggemann, N., Brunner, L., Cheedela, S. K., Das, S., Denissen, J., Dragaud, I., Dziekan, P., Ekblom, M., Engels, J. F., Esch, M., Forbes, R., Frauen, C., Freischem, L., García-Maroto, D., Geier, P., Gierz, P., González-Cervera, , Grayson, K., Griffith, M., Gutjahr, O., Haak, H., Hadade, I., Haslehner, K., ul Hasson, S., Hegewald, J., Kluft, L., Koldunov, A., Koldunov, N., Kölling, T., Koseki, S., Kosukhin, S., Kousal, J., Kuma, P., Kumar, A. U., Li, R., Maury, N., Meindl, M., Milinski, S., Mogensen, K., Niraula, B., Nowak, J., Praturi, D. S., Proske, U., Putrasahan, D., Redler, R., Santuy, D., Sármány, D., Schnur, R., Scholz, P., Sidorenko, D., Spät, D., Sützl, B., Takasuka, D., Tompkins, A., Uribe, A., Valentini, M., Veerman, M., Voigt, A., Warnau, S., Wachsmann, F., Wacławczyk, M., Wedi, N., Wieners, K.-H., Wille, J., Winkler, M., Wu, Y., Ziemen, F., Zimmermann, J., Bender, F. A.-M., Bojovic, D., Bony, S., Bordoni, S., Brehmer, P., Dengler, M., Dutra, E., Faye, S., Fischer, E., van Heerwaarden, C., Hohenegger, C., Järvinen, H., Jochum, M., Jung, T., Jungclaus, J. H., Keenlyside, N. S., Klocke, D., Konow, H., Klose, M., Malinowski, S., Martius, O., Mauritsen, T., Mellado, J. P., Mieslinger, T., Mohino, E., Pawłowska, H., Peters-von Gehlen, K.,

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Sarré, A., Sobhani, P., Stier, P., Tuppi, L., Vidale, P. L., Sandu, I., and Stevens, B.: nextGEMS: entering the era of kilometer-scale Earth system modeling, Geoscientific Model Development, 18, 7735–7761, https://doi.org/10.5194/gmd-18-7735-2025, 2025.

Sengupta, A., King, A. D., and Brown, J. R.: Do CMIP6 Models Capture Seasonal and Regional Differences in the Asymmetry of ENSO-Precipitation Teleconnections?, Journal of Geophysical Research: Atmospheres, 130, e2024JD041 031, https://doi.org/10.1029/2024JD041031, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2024JD041031, 2025.

Shen, Z., Duan, A., Li, D., and Li, J.: Assessment and Ranking of Climate Models in Arctic Sea Ice Cover Simulation: From CMIP5 to CMIP6, Journal of Climate, 34, 3609–3627, https://doi.org/10.1175/JCLI-D-20-0294.1, 2021.

Shepherd, T. G.: Atmospheric circulation as a source of uncertainty in climate change projections, Nature Geoscience, 7, 703–708, https://doi.org/10.1038/ngeo2253, 2014.

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, e2019RG000 678, https://doi.org/https://doi.org/10.1029/2019RG000678, tex.eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019RG000678, 2020.

Si, W., Liu, H., Zhang, X., and Zhang, M.: Double Intertropical Convergence Zones in Coupled Ocean-Atmosphere Models: Progress in CMIP6, Geophysical Research Letters, 48, https://doi.org/10.1029/2021GL094779, 2021.

Simpson, I. R., Shaw, T. A., Ceppi, P., Clement, A. C., Fischer, E., Grise, K. M., Pendergrass, A. G., Screen, J. A., Wills, R. C. J., Woollings, T., Blackport, R., Kang, J. M., and Po-Chedley, S.: Confronting Earth System Model trends with observations, Science Advances, 11, eadt8035, https://doi.org/10.1126/sciadv.adt8035, 2025.

Smith, S. C. and Ubilava, D.: The El Niño Southern Oscillation and economic growth in the developing world, Global Environmental Change, 45, 151–164, https://doi.org/10.1016/j.gloenvcha.2017.05.007, 2017.

Sohail, T., Irving, D. B., Zika, J. D., Holmes, R. M., and Church, J. A.: Fifty year trends in global ocean heat content traced to surface heat fluxes in the sub-polar ocean, Geophysical Research Letters, 48, e2020GL091 439, https://doi.org/https://doi.org/10.1029/2020GL091439, tex.eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020GL091439, 2021.

Song, X., Wang, D.-Y., Li, F., and Zeng, X.-D.: Evaluating the performance of CMIP6 Earth system models in simulating global vegetation structure and distribution, Advances in Climate Change Research, 12, 584–595, https://doi.org/10.1016/j.accre.2021.06.008, 2021.

Song, Z., Liu, H., and Chen, X.: Eastern equatorial Pacific SST seasonal cycle in global climate models: from CMIP5 to CMIP6, Acta Oceanologica Sinica, 39, 50–60, https://doi.org/10.1007/s13131-020-1623-z, 2020.

Spafford, L. and MacDougall, A. H.: Validation of terrestrial biogeochemistry in CMIP6 Earth system models: a review, Geoscientific Model Development, 14, 5863–5889, https://doi.org/10.5194/gmd-14-5863-2021, 2021.

Stephens, G. L.: Cloud feedbacks in the climate system: A critical review, Journal of climate, 18, 237–273, 2005.

Stouffer, R., Eyring, V., Meehl, G., Bony, S., Senior, C., Stevens, B., and Taylor, K.: CMIP5 scientific gaps and recommendations for CMIP6, Bulletin of the American Meteorological Society, 98, 95–105, https://doi.org/10.1175/BAMS-D-15-00013.1, 2017.

Sun, C. and Liang, X.-Z.: Understanding and Reducing Warm and Dry Summer Biases in the Central United States: Analytical Modeling to Identify the Mechanisms for CMIP Ensemble Error Spread, Journal of Climate, 36, 2035–2054, https://doi.org/10.1175/JCLI-D-22-0255.1, 2023.

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Sutton, R. T.: Climate science needs to take risk assessment much more seriously, Bulletin of the American Meteorological Society, 100, 1637 – 1642, https://doi.org/10.1175/BAMS-D-18-0280.1, 2019.

Swaminathan, R., Schewe, J., Walton, J., Zimmermann, K., Jones, C., Betts, R. A., Burton, C., Jones, C. D., Mengel, M., Reyer, C. P. O., Turner, A. G., and Weigel, K.: Regional impacts poorly constrained by climate sensitivity, Earth's Future, 12, e2024EF004 901, https://doi.org/https://doi.org/10.1029/2024EF004901, tex.eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2024EF004901, 2024.

Swingedouw, D., Houssais, M.-N., Herbaut, C., Blaizot, A.-C., Devilliers, M., and Deshayes, J.: AMOC Recent and Future Trends: A Crucial Role for Oceanic Resolution and Greenland Melting?, Frontiers in Climate, 4, https://doi.org/10.3389/fclim.2022.838310, 2022.

Takano, Y., Ilyina, T., Tjiputra, J., Eddebbar, Y., Berthet, S., Bopp, L., Buitenhuis, E., Butenschön, M., Christian, J. R., Dunne, J. P., Gröger, M., Hayashida, H., Hieronymus, J., Koenigk, T., Krasting, J. P., Long, M. C., Lovato, T., Nakano, H., Palmieri, J., Schwinger, J., Séférian, R., Suntharalingam, P., Tatebe, H., Tsujino, H., Urakawa, S., Watanabe, M., and Yool, A.: Simulations of ocean deoxygenation in the historical era: insights from forced and coupled models, Frontiers in Marine Science, 10, 1139 917, https://doi.org/10.3389/fmars.2023.1139917, 2023.

Talib, J., Müller, O. V., Barton, E. J., Taylor, C. M., and Vidale, P. L.: The Representation of Soil Moisture-Atmosphere Feedbacks across the Tibetan Plateau in CMIP6, Advances in Atmospheric Sciences, 40, 2063–2081, https://doi.org/10.1007/s00376-023-2296-2, 2023.

Taschetto, A. S., Ummenhofer, C. C., Stuecker, M. F., Dommenget, D., Ashok, K., Rodrigues, R. R., and Yeh, S.-W.: ENSO Atmospheric Teleconnections, in: El Niño Southern Oscillation in a Changing Climate, pp. 309–335, American Geophysical Union (AGU), ISBN 978-1-119-54816-4, https://doi.org/10.1002/9781119548164.ch14, section: 14 _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/9781119548164.ch14, 2020.

Tian, B. and Dong, X.: The Double-ITCZ Bias in CMIP3, CMIP5, and CMIP6 Models Based on Annual Mean Precipitation, Geophysical Research Letters, 47, e2020GL087 232, https://doi.org/10.1029/2020GL087232, 2020.

Tian, B., Fetzer, E. J., and Teixeira, J.: Assessing the Tropospheric Temperature and Humidity Simulations in CMIP3/5/6 Models Using the AIRS Obs4MIPs V2.1 Data, Journal of Geophysical Research: Atmospheres, 129, e2023JD040 536, https://doi.org/10.1029/2023JD040536, 2024.

Tjiputra, J. F., Negrel, J., and Olsen, A.: Early detection of anthropogenic climate change signals in the ocean interior, Scientific Reports, 13, 3006, https://doi.org/10.1038/s41598-023-30159-0, 2023.

Todd-Brown, K. E. O., Randerson, J. T., Post, W. M., Hoffman, F. M., Tarnocai, C., Schuur, E. a. G., and Allison, S. D.: Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations, Biogeosciences, 10, 1717–1736, https://doi.org/10.5194/bg-10-1717-2013, 2013.

Trenberth, K. E.: ENSO in the Global Climate System, in: El Niño Southern Oscillation in a Changing Climate, pp. 21–37, American Geophysical Union (AGU), ISBN 978-1-119-54816-4, https://doi.org/10.1002/9781119548164.ch2, section: 2 _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/9781119548164.ch2, 2020.

Trivedi, S., Hobbs, W., and Raphael, M.: An assessment of Antarctic sea-ice thickness in CMIP6 simulations with comparison to the satellite-based observations and reanalyses, The Cryosphere, 19, 6771–6790, https://doi.org/10.5194/tc-19-6771-2025, 2025.

Turetsky, M. R., Abbott, B. W., Jones, M. C., Anthony, K. W., Olefeldt, D., Schuur, E. A. G., Grosse, G., Kuhry, P., Hugelius, G., Koven, C., Lawrence, D. M., Gibson, C., Sannel, A. B. K., and McGuire, A. D.: Carbon release through abrupt permafrost thaw, Nature Geoscience, 13, 138–143, https://doi.org/10.1038/s41561-019-0526-0, 2020.

Vaittinada Ayar, P., Bopp, L., Christian, J. R., Ilyina, T., Krasting, J. P., Séférian, R., Tsujino, H., Watanabe, M., Yool, A., and Tjiputra, J.: Contrasting projections of the ENSO-driven CO_2 flux variability in the equatorial Pacific under high-warming scenario, Earth System Dynamics, 13, 1097–1118, https://doi.org/10.5194/esd-13-1097-2022, 2022.

van Westen, R. M. and Dijkstra, H. A.: Persistent climate model biases in the Atlantic Ocean's freshwater transport, Ocean Science, 20, 549–567, https://doi.org/10.5194/os-20-549-2024, 2024.

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

Varney, R. M., Chadburn, S. E., Friedlingstein, P., Burke, E. J., Koven, C. D., Hugelius, G., and Cox, P. M.: A spatial emergent constraint on the sensitivity of soil carbon turnover to global warming, Nature Communications, 11, 5544, https://doi.org/10.1038/s41467-020-19208-8, 2020.

Varney, R. M., Chadburn, S. E., Burke, E. J., and Cox, P. M.: Evaluation of soil carbon simulation in CMIP6 Earth system models, Biogeosciences, 19, 4671–4704, https://doi.org/10.5194/bg-19-4671-2022, 2022.

Varney, R. M., Chadburn, S. E., Burke, E. J., Jones, S., Wiltshire, A. J., and Cox, P. M.: Simulated responses of soil carbon to climate change in CMIP6 Earth system models: the role of false priming, Biogeosciences, 20, 3767–3790, https://doi.org/10.5194/bg-20-3767-2023, 2023.

Varney, R. M., Friedlingstein, P., Chadburn, S. E., Burke, E. J., and Cox, P. M.: Soil carbon-concentration and carbon-climate feedbacks in CMIP6 Earth system models, Biogeosciences, 21, 2759–2776, https://doi.org/10.5194/bg-21-2759-2024, 2024.

Vial, J., Dufresne, J.-L., and Bony, S.: On the interpretation of inter-model spread in CMIP5 climate sensitivity estimates, Climate Dynamics, 41, 3339–3362, https://doi.org/10.1007/s00382-013-1725-9, 2013.

Vijayeta, A. and Dommenget, D.: An evaluation of ENSO dynamics in CMIP simulations in the framework of the recharge oscillator model, Climate Dynamics, 51, 1753–1771, https://doi.org/10.1007/s00382-017-3981-6, 2018.

Wang, A., Kong, X., Chen, Y., and Ma, X.: Evaluation of Soil Moisture in CMIP6 Multimodel Simulations Over Conterminous China, Journal of Geophysical Research: Atmospheres, 127, e2022JD037 072, https://doi.org/10.1029/2022JD037072, 2022a.

Wang, B., Jin, C., and Liu, J.: Understanding Future Change of Global Monsoons Projected by CMIP6 Models, Journal of Climate, 33, 6471–6489, https://doi.org/10.1175/JCLI-D-19-0993.1, 2020.

Wang, C., Zhang, L., Lee, S.-K., Wu, L., and Mechoso, C. R.: A global perspective on CMIP5 climate model biases, Nature Climate Change, 4, 201–205, 2014.

Wang, C., Zou, L., and Zhou, T.: SST biases over the Northwest Pacific and possible causes in CMIP5 models, Science China Earth Sciences, 61, 792–803, 2018.

Wang, J., DeFlorio, M. J., Kim, H., Guirguis, K., and Gershunov, A.: Future MJO Change and Its Impact on Extreme Precipitation and Temperature Over the Western US in CMIP6, Journal of Geophysical Research: Atmospheres, 130, e2024JD042 123, https://doi.org/10.1029/2024JD042123, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2024JD042123, 2025.

Wang, S., Wang, Q., Wang, M., Lohmann, G., and Qiao, F.: Arctic Ocean Freshwater in CMIP6 Coupled Models, Earth's Future, 10, e2022EF002 878, https://doi.org/10.1029/2022EF002878, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2022EF002878, 2022b.

Watts, M., Maslowski, W., Lee, Y. J., Kinney, J. C., and Osinski, R.: A Spatial Evaluation of Arctic Sea Ice and Regional Limitations in CMIP6 Historical Simulations, Journal of Climate, 34, 6399–6420, https://doi.org/10.1175/JCLI-D-20-0491.1, 2021.

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Weijer, W., Cheng, W., Garuba, O. A., Hu, A., and Nadiga, B. T.: CMIP6 Models Predict Significant 21st Century Decline of the Atlantic Meridional Overturning Circulation, Geophysical Research Letters, 47, e2019GL086 075, https://doi.org/10.1029/2019GL086075, 2020.

Weijer, W., Huo, Y., Garuba, O., Lee, Y., Molodtsov, S., and Kurtakoti, P.: What CMIP6 Models Tell Us About the Impact of AMOC Variability on the Arctic, Geophysical Research Letters, 52, e2025GL116 282, https://doi.org/10.1029/2025GL116282, 2025.

Wild, M.: The global energy balance as represented in CMIP6 climate models, Climate Dynamics, 55, 553–577, 2020.

Wild, M., Folini, D., Hakuba, M. Z., Schär, C., Seneviratne, S. I., Kato, S., Rutan, D., Ammann, C., Wood, E. F., and König-Langlo, G.: The energy balance over land and oceans: an assessment based on direct observations and CMIP5 climate models, Climate Dynamics, 44, 3393–3429, 2015.

Wild, M., Hakuba, M. Z., Folini, D., Dörig-Ott, P., Schär, C., Kato, S., and Long, C. N.: The cloud-free global energy balance and inferred cloud radiative effects: an assessment based on direct observations and climate models, Climate Dynamics, 52, 4787–4812, 2019.

Winkelmann, R., Dennis, D. P., Donges, J. F., Loriani, S., Klose, A. K., Abrams, J. F., Alvarez-Solas, J., Albrecht, T., Armstrong McKay, D., Bathiany, S., Blasco Navarro, J., Brovkin, V., Burke, E., Danabasoglu, G., Donner, R. V., Drüke, M., Georgievski, G., Goelzer, H., Harper, A. B., Hegerl, G., Hirota, M., Hu, A., Jackson, L. C., Jones, C., Kim, H., Koenigk, T., Lawrence, P., Lenton, T. M., Liddy, H., Licón-Saláiz, J., Menthon, M., Montoya, M., Nitzbon, J., Nowicki, S., Otto-Bliesner, B., Pausata, F., Rahmstorf, S., Ramin, K., Robinson, A., Rockström, J., Romanou, A., Sakschewski, B., Schädel, C., Sherwood, S., Smith, R. S., Steinert, N. J., Swingedouw, D., Willeit, M., Weijer, W., Wood, R., Wyser, K., and Yang, S.: The Tipping Points Modelling Intercomparison Project (TIPMIP): Assessing tipping point risks in the Earth system, EGUsphere, pp. 1–52, https://doi.org/10.5194/egusphere-2025-1899, 2025.

Woollings, T., Barriopedro, D., Methven, J., Son, S.-W., Martius, O., Harvey, B., Sillmann, J., Lupo, A. R., and Seneviratne, S.: Blocking and its Response to Climate Change, Current Climate Change Reports, 4, 287–300, https://doi.org/10.1007/s40641-018-0108-z, 2018.

Wunderling, N., von der Heydt, A. S., Aksenov, Y., Barker, S., Bastiaansen, R., Brovkin, V., Brunetti, M., Couplet, V., Kleinen, T., Lear, C. H., Lohmann, J., Roman-Cuesta, R. M., Sinet, S., Swingedouw, D., Winkelmann, R., Anand, P., Barichivich, J., Bathiany, S., Baudena, M., Bruun, J. T., Chiessi, C. M., Coxall, H. K., Docquier, D., Donges, J. F., Falkena, S. K. J., Klose, A. K., Obura, D., Rocha, J., Rynders, S., Steinert, N. J., and Willeit, M.: Climate tipping point interactions and cascades: a review, Earth System Dynamics, 15, 41–74, https://doi.org/10.5194/esd-15-41-2024, 2024.

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

Xiang, B., Wang, B., Ding, Q., Jin, F., Fu, X., and Kim, H.-J.: Reduction of the thermocline feedback associated with mean SST bias in ENSO simulation, Climate Dynamics, 39, 1413–1430, https://doi.org/10.1007/s00382-011-1164-4, 2012.

Xu, W., Zheng, X.-T., and Hu, K.: Revisiting the modelled cloud radiative feedback on ENSO — the source of model uncertainty, Climate Dynamics, 63, 237, https://doi.org/10.1007/s00382-025-07728-2, 2025.

Yan, X., Zhang, R., and Knutson, T. R.: Underestimated AMOC Variability and Implications for AMV and Predictability in CMIP Models, Geophysical Research Letters, 45, 4319–4328, https://doi.org/10.1029/2018GL077378, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2018GL077378, 2018.

Yan, Y., Zhu, C., Liu, B., and Jiang, S.: Annual Cycle of East Asian Precipitation Simulated by CMIP6 Models, Atmosphere, 12, 24, https://doi.org/10.3390/atmos12010024, 2020.

Yeager, S. G. and Robson, J. I.: Recent Progress in Understanding and Predicting Atlantic Decadal Climate Variability, Current Climate Change Reports, 3, 112–127, https://doi.org/10.1007/s40641-017-0064-z, 2017.

https://doi.org/10.5194/egusphere-2026-3510
Preprint. Discussion started: 16 July 2026

Yu, L., Zhong, S., Sui, C., and Sun, B.: A regional and seasonal approach to explain the observed trends in the Antarctic sea ice in recent decades, International Journal of Climatology, 43, 2953–2974, https://doi.org/https://doi.org/10.1002/joc.8010, 2023.

Yuan, K., Zhu, Q., Riley, W. J., Li, F., and Wu, H.: Understanding and reducing the uncertainties of land surface energy flux partitioning within CMIP6 land models, Agricultural and Forest Meteorology, 319, 108 920, https://doi.org/10.1016/j.agrformet.2022.108920, 2022.

Yuan, S., Quiring, S. M., and Leasor, Z. T.: Historical Changes in Surface Soil Moisture Over the Contiguous United States: An Assessment of CMIP6, Geophysical Research Letters, 48, e2020GL089 991, https://doi.org/10.1029/2020GL089991, 2021.

Yuval, J. and O'Gorman, P. A.: Stable machine-learning parameterization of subgrid processes for climate modeling at a range of resolutions, Nature Communications, 11, 3295, https://doi.org/10.1038/s41467-020-17142-3, 2020.

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, e2019GL085 782, https://doi.org/10.1029/2019GL085782, 2020.

Zhang, G., Feng, M., Chen, M., Li, M., and Wang, D.: Influences of meridional SST gradient biases on subsurface temperature in the tropical indian ocean: Roles of SMOC and ITF, Geophysical Research Letters, 53, e2025GL118 928, 2026.

Zhang, L. and Zhao, C.: Processes and mechanisms for the model sst biases in the n orth a tlantic and n orth p acific: a link with the a tlantic meridional overturning circulation, Journal of Advances in Modeling Earth Systems, 7, 739–758, 2015.

Zhang, Q., Liu, B., Li, S., and Zhou, T.: Understanding Models' Global Sea Surface Temperature Bias in Mean State: From CMIP5 to CMIP6, Geophysical Research Letters, 50, e2022GL100 888, https://doi.org/10.1029/2022GL100888, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2022GL100888, 2023.

Zhang, S., Hu, Y., Zhang, J., and Xia, Y.: Attribution of Biases of Interhemispheric Temperature Contrast in CMIP6 Models, Advances in Atmospheric Sciences, 41, 325–340, https://doi.org/10.1007/s00376-023-3002-0, 2024.

Zhao, A., Robson, J., Sutton, R., Lai, M. W., Mecking, J. V., Yeager, S., and Petit, T.: Large diversity in AMOC internal variability across NEMO-based climate models, Climate Dynamics, 62, 3355–3374, https://doi.org/10.1007/s00382-023-07069-y, 2024a.

Zhao, X., Liu, X., Lin, L., Qin, Y., Zelinka, M., Klein, S., Zhang, M., Zhang, K., Ma, P.-L., Zhu, J., Lu, Z., and Saravanan, R.: Larger Cloud Liquid Water Enhances Both Aerosol Indirect Forcing and Cloud Radiative Feedback in Two Earth System Models, Geophysical Research Letters, 51, https://doi.org/10.1029/2023GL105529, 2024b.

Zheng, X., Tao, C., Zhang, C., Xie, S., Zhang, Y., Xi, B., and Dong, X.: Assessment of CMIP5 and CMIP6 AMIP Simulated Clouds and Surface Shortwave Radiation Using ARM Observations over Different Climate Regions, Journal of Climate, 36, 8475–8495, https://doi.org/10.1175/JCLI-D-23-0247.1, 2023.

Zheng, Y., Shinoda, T., Lin, J.-L., and Kiladis, G. N.: Sea surface temperature biases under the stratus cloud deck in the southeast Pacific Ocean in 19 IPCC AR4 coupled general circulation models, Journal of Climate, 24, 4139–4164, 2011.

Zhou, S., Yu, B., Lintner, B. R., Findell, K. L., and Zhang, Y.: Projected increase in global runoff dominated by land surface changes, Nature Climate Change, 13, 442–449, https://doi.org/https://doi.org/10.1038/s41558-023-01659-8, 2023.

Zhu, J., Poulsen, C. J., and Otto-Bliesner, B. L.: High climate sensitivity in CMIP6 model not supported by paleoclimate, Nature Climate Change, 10, 378–379, 2020.

Zickfeld, K., MacIsaac, A. J., Canadell, J. G., Fuss, S., Jackson, R. B., Jones, C. D., Lohila, A., Matthews, H. D., Peters, G. P., Rogelj, J., and Zaehle, S.: Net-zero approaches must consider Earth system impacts to achieve climate goals, Nature Climate Change, 13, 1298–1305, https://doi.org/10.1038/s41558-023-01862-7, 2023.