

UC Irvine

UC Irvine Previously Published Works

Title

Radiative, Hydrologic, and Circulation Responses to Warming in Cess-Potter Simulations Using the Global 3.25-km SCREAM

Permalink

<https://escholarship.org/uc/item/5t65r0x2>

Journal

Journal of Advances in Modeling Earth Systems, 17(10)

ISSN

1942-2466

Authors

Terai, CR

Keen, ND

Caldwell, PM

et al.

Publication Date

2025-10-01

DOI

10.1029/2024ms004930

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed



RESEARCH ARTICLE

10.1029/2024MS004930

Special Collection:

The Energy Exascale Earth
System Model (E3SM):
Advancements in a Decade of
Earth System Modeling

Key Points:

- We run a pair of 13-month simulations to quantify the atmospheric climate feedbacks and their breakdown in global storm-resolving SCREAM 3 km
- SCREAM 3 km's climate feedback is at the high end of climate models and suggests that km-scale models may not narrow the spread in feedbacks
- More negative cloud feedbacks in SCREAM 12 km than in SCREAM 3 km imply resolution sensitivities in feedbacks even in sub-20 km global models

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

C. R. Terai,
terai1@lln.gov

Citation:

Terai, C. R., Keen, N. D., Caldwell, P. M., Beydoun, H., Bogenschutz, P. A., Chao, L.-W., et al. (2025). Radiative, hydrologic, and circulation responses to warming in Cess-Potter simulations using the global 3.25-km SCREAM. *Journal of Advances in Modeling Earth Systems*, 17, e2024MS004930. <https://doi.org/10.1029/2024MS004930>

Received 31 DEC 2024

Accepted 25 SEP 2025

© 2025 Battelle Memorial Institute and The Author(s). Journal of Advances in Modeling Earth Systems published by Wiley Periodicals LLC on behalf of American Geophysical Union.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Radiative, Hydrologic, and Circulation Responses to Warming in Cess-Potter Simulations Using the Global 3.25-km SCREAM

C. R. Terai¹ , N. D. Keen², P. M. Caldwell¹ , H. Beydoun¹ , P. A. Bogenschutz¹ , L.-W. Chao¹ , B. R. Hillman³ , H.-Y. Ma¹, M. D. Zelinka¹ , L. Bertagna³, A. M. Bradley³, T. C. Clevenger³, A. S. Donahue¹ , J. Foucar³ , J.-C. Golaz¹ , O. Guba³ , W. Hannah¹ , J. Lee¹ , W. Lin⁴, N. Mahfouz⁵ , J. Mülmenstädt⁵, A. G. Salinger³ , B. Singh⁵ , S. Sreepathi⁶ , Y. Qin⁵ , M. A. Taylor³ , P. A. Ullrich¹ , W.-Y. Wu¹ , X. Yuan⁷, C. S. Zender⁸ , and Y. Zhang¹ 

¹Lawrence Livermore National Laboratory, Livermore, CA, USA, ²Lawrence Berkeley National Laboratory, Berkeley, CA, USA, ³Sandia National Laboratories, Albuquerque, NM, USA, ⁴Brookhaven National Laboratory, Upton, NY, USA, ⁵Pacific Northwest National Laboratory, Richland, WA, USA, ⁶Oak Ridge National Laboratory, Oak Ridge, TN, USA, ⁷Argonne National Lab, Lemont, IL, USA, ⁸University of California, Irvine, Irvine, CA, USA

Abstract Using the global 3.25-km Simple Cloud Resolving E3SM Atmosphere Model (SCREAM 3 km), a pair of 13-month Cess-Potter simulations are performed to quantify the radiative feedbacks and the hydrologic and circulation responses to warming. Large-scale aspects of SCREAM 3 km's top-of-atmosphere radiative fluxes, precipitation rates, and circulations are in good agreement with observations and reanalysis, with notable differences, including a drier lower free-troposphere in the Tropics, reduced precipitation and humidity over the Tropical West Pacific, and poleward shifted Southern Hemisphere midlatitude jet. In response to warming, SCREAM 3 km predicts a total radiative feedback within the top 15% of the CMIP5 and CMIP6 models, which puts it substantially higher than the feedback reported by other kilometer-scale models. SCREAM 3 km's high radiative feedback stems from a strongly positive shortwave cloud feedback, most prominent over the mid- and high-latitudes. SCREAM 3 km's high precipitation response also puts it among the highest of CMIP models, whereas its circulation response is within the spread of CMIP models. An ensemble of five perturbed initial condition Cess-Potter simulations with a 12 km version of SCREAM (SCREAM 12 km) is performed to characterize uncertainty and resolution sensitivity. It suggests that the uncertainty from analyzing a pair of 1-year simulations is small compared to the inter-model spread in feedbacks and precipitation response. SCREAM 12 km also produces a strong precipitation response to warming but a much lower cloud feedback and total radiative feedback. The results from these experiments suggest that the spread in climate feedbacks will likely persist in the next generation of kilometer-scale models.

Plain Language Summary To determine how the large-scale atmosphere responds to warming, we perform a pair of 13-month simulations using the global 3.25-km Simple Cloud Resolving E3SM Atmosphere Model (SCREAM 3 km), where one simulation uses present-day sea-surface temperatures and the other uses sea-surface temperatures that are increased uniformly by 4 K. SCREAM 3 km shows relatively good skill in representing cloud properties, precipitation patterns, and large-scale wind patterns that we observe in nature, but also simulates an overly dry lower atmosphere over the Tropics and a western Pacific Ocean region that lacks precipitation. Compared to current-generation climate models with coarser ~100 km grid resolutions, SCREAM 3 km produces a climate feedback that is on the high-end of current-generation models, which suggests a greater warming in response to increased greenhouse gas levels. SCREAM 3 km's higher climate feedback mainly stems from decreasing cloud reflection over the subtropical and mid-latitude oceans. On the other hand, similar experiments performed with a 12 km version of SCREAM produces climate feedbacks close to the average feedback of current-generation models. The results from these experiments suggest that the spread in climate feedbacks will likely persist in the next generation of kilometer-scale models.

1. Introduction

More than 35 years ago, Cess and Potter (1988) proposed and reported on a computationally efficient set of simulations to predict the radiative feedbacks from atmosphere-land-ocean coupled simulations with perturbed

carbon dioxide concentrations. The experiments consisted of a pair of prescribed sea surface temperature (SST) simulations where the SSTs are uniformly perturbed by $\pm 2\text{K}$. A follow-on study using an ensemble of state-of-the-art climate models at the time found that most of the intermodel spread in the total radiative feedback stemmed from variations in their predictions of how clouds respond to warming (Cess et al., 1989). Even in more recent generations of climate model ensembles, these more constrained and simplified atmosphere-only simulations, which we refer to as Cess-Potter experiments, continue to be useful to predict feedbacks from fully coupled experiments and useful to better understand cloud feedbacks (e.g., Qin et al., 2022; Ringer et al., 2014; Sherwood et al., 2014; Webb et al., 2015, 2024).

Despite years of work to better characterize, understand, and constrain cloud feedbacks, the spread in cloud feedbacks continues to be large and remains the largest source of model uncertainty in estimating the warming from increased carbon dioxide concentrations (Sherwood et al., 2020; Zelinka et al., 2020). In this context, some studies have associated the spread in cloud feedbacks with our reliance on parameterizations of convection (e.g., Sherwood et al., 2014). The parameterizations of convection, which represent the precipitation and vertical mixing processes of unresolved convective clouds, are necessary in climate models that have typical horizontal resolutions of approximately 100 km. Their potential impacts are not only limited to deep convective clouds and cirrus clouds that detrain from them, but also include low-level clouds, through the parameterizations' impact on boundary-layer-free-troposphere mixing (Qin et al., 2024; Sherwood et al., 2014; Zhang & Bretherton, 2008), their impacts on high-level cloud cover that affect cloud-top radiative cooling (Christensen et al., 2013), and their remote impact on the temperature and humidity structures of the free-tropospheric air overlying low-clouds (Brient & Bony, 2013; Schiro et al., 2022; Zhou et al., 2017).

In recent years, computational advancements have allowed a growing number of modeling centers to develop and perform simulations with kilometer-scale global storm-resolving models (GSRMs), with horizontal grid spacings of less than 5 km, such that two phases of model intercomparisons of 40-day simulations (Dynamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains—DYAMOND) had participation from almost a dozen modeling centers (Stevens et al., 2019). The Simple Cloud Resolving E3SM Atmosphere Model (SCREAM) is one such GSRM developed under the US Department of Energy, which participated in the second phase of DYAMOND (Caldwell et al., 2021) and recently ran four 40-day simulations (Donahue et al., 2024). Performing a pair of multi-decadal simulations following the protocol of the Atmosphere Model Intercomparison Project (AMIP) with GSRMs is still some years away with current computational resources. Qin et al. (2022) reported that while a 1-year AMIP-type simulation is sufficient to capture the “inter-model spread” of cloud feedbacks in coupled simulations, longer simulations provide more accurate estimates for individual models. However, for the purposes of this study, we adopt the 1-year simulation approach based on its demonstrated capability, as later analyzed using ensemble members of SCREAM 12 km. Furthermore, Merlis et al. (2024) recently reported total radiative feedback estimates from the kilometer-scale X-SHIELD model from running a pair of 2-year simulations. With these recent results, we are well-poised to start comparing, understanding, and quantifying the atmospheric climate feedbacks coming from this new generation of models.

This paper aims to provide an overview of the Cess-Potter experiments performed using the SCREAMv1 model described in Donahue et al. (2024). The simulations, performed using the default 3.25 km horizontal resolution (SCREAM 3 km) and a coarser 12 km resolution (SCREAM 12 km) will provide some insight into whether atmospheric climate feedbacks differ from those predicted by climate models with ~ 100 km resolution that are part of CMIP and by the handful of existing km-scale simulations. We provide a description of the model and the experimental design in Section 2, report on the large-scale atmospheric features in the baseline SCREAM 3 km simulation in Section 3, compare and contrast the climate feedbacks against CMIP-class models in Section 4, examine some of the processes underlying feedback differences in Section 5, explore the hydrologic and circulation response in Section 6, describe how mesoscale convective systems in SCREAM 3 km respond to warming in Section 7, and finally discuss the implications of these results and provide a glimpse of future inquiry made possible with the model experiments at this resolution in Section 8.

2. Methods

The main experiments analyzed in this study are a pair of Cess-Potter experiments performed using the 3.25-km SCREAM simulations. Details of the model are found in Section 2.1, where details of the experiment design can be found in Section 2.2. We also performed five pairs of Cess-Potter experiments using a coarser 12-km version of

SCREAM to examine the uncertainties arising from internal variability; their details can be found in Section 2.3. Finally, details of the observations, reanalysis, and CMIP models that we compare the SCREAM results against are found in Section 2.4, and the calculation of radiative feedbacks in this study are described in Section 2.5.

2.1. SCREAMv1 and Machines Used

As described in Donahue et al. (2024), SCREAMv1 is a new GSRM written in C++ and employing the Kokkos C++ Performance Portability Ecosystem (Trott et al., 2022) to enable it to run efficiently on a variety of hardware architectures, including conventional CPU-based systems and emerging GPU-based systems (from multiple vendors). For this work we use the Frontier exascale supercomputer at OLCF with 2048 nodes for 3 km experiments and 96 nodes for 12 km experiments.

The backbone of SCREAMv1 is the non-hydrostatic version of the High Order Method Modeling Environment (HOMME) dynamical core used in all E3SM model releases (Dennis et al., 2005, 2012; Taylor et al., 2020), which was written in C++ for SCREAMv1 (Bertagna et al., 2020). Turbulence, condensation, and liquid cloud fraction is handled by the Simplified Higher Order Closure (SHOC) scheme (Bogenschütz & Krueger, 2013). All-or-nothing ice cloud fraction is assumed with an ice mixing ratio threshold of 10^{-14} kg kg⁻¹. The single-category Predicted Particle Properties (P3) of Morrison and Milbrandt (2015) is used for microphysics. As described in Caldwell et al. (2021), liquid supersaturation was removed from P3 to enable consistency with SHOC's liquid saturation adjustment assumption. Radiation is handled by a C++ version of the RTE + RRTMGP radiative transfer package originally written in Fortran (Donahue et al., 2024; Pincus et al., 2019). This set of simulations also uses the Simplified Prescribed Aerosol (SPA) scheme to prescribe the cloud droplet number concentration and aerosol radiative properties coming from a previous E3SM simulation, as described by Donahue et al. (2024) and Mahfouz et al. (2025). The version of the SPA data used in these experiments can be found at Terai (2024b).

The model closely follows what was reported by Donahue et al. (2024) with the following changes to address specific model biases. First, the Turbulent Mountain Stress scheme was added to improve near-surface wind speeds. The University of Arizona surface flux scheme (Reeves Eyre et al., 2021) is also used over the existing E3SMv1 scheme to reduce the high wind speed and high latent heat flux bias over the oceans. Second, improvements were made to SHOC to address nocturnal over-mixing, which had previously led to warmer 2-m temperatures. Third, cloud-related parameters were tuned in SHOC and in P3 to increase low-level clouds and to increase the upper threshold for ice number concentration, which has the effect of increasing high-level clouds (Donahue et al., 2024). Finally, improvements to the water conservation were made such that the water is conserved to within 34 mm/century in these experiments.

2.2. Experimental Design

We run a pair of 13-month simulations, one using present-day forcings and sea-surface temperatures (SSTs) and the other where the SSTs are increased uniformly by 4K (+4K simulation). Although often referred to as atmosphere-only simulations, the interactive E3SM Land Model is turned on in these simulations, so they are technically atmosphere-land coupled simulations. Running with an interactive land allows for land-atmosphere interactions to occur that affect both the present-day simulations and their response to warmed SSTs. The simulations span from 1 Aug 2019 to 31 Aug 2020, with analysis performed on the last 12 months of the simulation. The sea-surface temperature and sea-ice coverage are taken from the Group for High Resolution Sea Surface Temperature (GHRSSST) Level 4 sea surface temperature analysis produced daily on an operational basis at the UK Met Office (Donlon et al., 2012).

The choice of simulation period is motivated by three aspects: First, the period does not exhibit strong SST anomalies associated with the El Niño Southern Oscillation (ENSO), which is important because Qin et al. (2022) noted that cloud feedbacks in Cess-Potter simulations show better agreement with coupled simulations during ENSO-neutral years. Second, a relatively recent period allows the evaluation against observations to use more recent satellite data sets (e.g., The Global Precipitation Measurement Mission). Finally, we wanted to start prior to the northern hemisphere winter to allow for a natural accumulation of continental snow pack, rather than relying on the snow pack to be provided by the initial condition. Although the simulation period does not match the March 2020 to February 2021 period proposed by Takasuka, Satoh, et al. (2024) for a year-long DYAMOND inter-comparison, six-months of the 12-month analysis period overlap, and the potential impacts on feedbacks are explored with SCREAM 12 km, as discussed below.

Table 1

Time Steps and Subcycling Frequencies Associated With SCREAM 3 km and SCREAM 12 km

Time steps/frequency	SCREAM 3 km	SCREAM 12 km
Physics time step (s)	100	600
Radiation time step (s)	300	1800
Dynamics time step (s)	8.333	33.333
Macrophysics/microphysics subcycle frequency	1	6

Note. Macrophysics-microphysics subcycle frequency indicates the number of times the macrophysics and microphysics schemes are subcycled within a physics time step.

The land initial condition was created by running the land model with prescribed atmospheric forcing from CRUNCEP Version 7 (Viovy, 2018). The land spin-up run is configured to stop when reaching the simulation start date, for a total of 28 and 20 years, respectively for the 3 and 12 km resolutions.

2.3. SCREAM 12 km Ensemble

In addition to the 3 km version of SCREAM, we also perform the same Cess-Potter experiments using a coarser 12 km version of SCREAM without any changes to the physics parameters. Time steps are increased to match the decrease in the resolution (Table 1), but everything else is kept the same. One benefit of the 12 km simulations is that their smaller computational cost allows for multiple simulations to investigate uncertainties associated with internal variability. On the other hand, we recognize that running at 12 km resolution without a deep convective scheme may lead to a less realistic representation of convection and hence clouds and precipitation. To produce the additional four-member initial condition ensemble simulations with SCREAM 12 km, we apply small random perturbations (<0.0001 K) to the initial temperature field so that simulations follow different weather trajectories. Furthermore, to assess the sensitivity of the warming response to the choice of time period, we extend one of the simulations out to 28 February 2021 and analyze the period of March 2020 to February 2021. This period matches the period proposed by Takasuka, Satoh, et al. (2024) for the year-long DYAMOND intercomparison.

A comprehensive list of model outputs from both sets of SCREAM 3 and 12 km simulations can be found in Table S2–S4 in Supporting Information S1.

2.4. CMIP Models and Observations

To provide context to the SCREAM 3 km results, we also plot results from the Cloud Feedback Model Intercomparison Project (CFMIP) using models that participated in the Atmosphere Model Intercomparison Project (AMIP) and the AMIP +4K protocols (Webb et al., 2017). A total of 24 models from CMIP5 and CMIP6 are used in the radiative feedback analysis (Sect. 4) and for the other portions, we compare the SCREAM 3 km response to 12 CMIP6 models. A list of the models used in this study can be found in Table S1 in Supporting Information S1.

For observational data sets to evaluate the present-day simulation, the monthly averaged satellite-based retrievals of Integrated Multi-satellite Retrievals for the Global Precipitation Measurement (GPM) Mission v7 (IMERG; Huffman, Bolvin, et al., 2020) are used for precipitation rate. Top-of-atmosphere radiative fluxes are used from edition 4.2 of the Clouds and the Earth's Radiant Energy System Energy Balanced and Filled Data Product (CERES-EBAF; Loeb et al., 2018; Kato et al., 2018). For height-resolved temperature, humidity, and winds we use the fifth-generation ECMWF atmospheric reanalysis (ERA5; Hersbach et al., 2020). For comparison with CMIP6 precipitation rates, precipitation retrievals from Global Precipitation Climatology Project (GPCP) v3.2 are used (Huffman, 2021; Huffman, Adler, et al., 2020). In comparisons with SCREAM, unless otherwise mentioned, the SCREAM outputs are conservatively regridded to the native grid of the corresponding observational data set.

2.5. Calculating the Radiative Feedbacks

The radiative feedback parameter λ is a useful quantity to compare the sensitivity of the global temperature response to an external radiative forcing from models and observational estimates alike. It quantifies the change in global-mean surface temperature ΔT_s that occurs in response to an instantaneous radiative perturbation to the

system F as the net top of atmosphere imbalance ΔN equilibrates to zero, and is expressed as its inverse such that its units are $\text{W m}^{-2} \text{K}^{-1}$ as in Equation 1:

$$\Delta N = F + \lambda \Delta T_s. \quad (1)$$

Various ways of estimating feedbacks exist, including performing slab-ocean experiments (Kiehl et al., 2006; Manabe & Stouffer, 1980), running atmosphere-ocean-coupled simulations over the historical period (Gregory & Andrews, 2016), or inferring from observational records (Sherwood et al., 2020). Often, feedbacks are calculated from atmosphere-ocean-coupled simulations where the atmospheric CO_2 concentrations are abruptly doubled or quadrupled from pre-industrial conditions. The rate at which the net top-of-atmosphere radiative imbalance returns to equilibrium (zero) as a function of the change in global-mean temperature gives the λ value (Gregory et al., 2004). However, as Cess and Potter (1988) suggested, and others have since verified (Qin et al., 2022; Ringer et al., 2014), the feedback parameter can also be calculated from a pair of atmosphere-only simulations forced by perturbed SSTs. In this scenario, F is zero, so the feedback λ can be calculated from the differences in the net top-of-atmosphere radiation N and surface temperature T_s between the two simulations ($\Delta N / \Delta T_s$). This total feedback can further be decomposed into individual components using the radiative kernel method (Soden et al., 2008). We use the method outlined by Qin et al. (2022) to calculate the feedbacks such that the feedbacks for CMIP models are those reported in their study. The differences in the monthly mean fields (temperature, water vapor, and albedo) between the two experiments are multiplied by the radiative kernels, and for atmospheric temperature and water vapor, integrated from the surface to the tropopause. That is then divided by the global mean difference in the surface temperature (ΔT_s) to obtain the feedbacks. We use the relative humidity as the basis for our decomposition, as proposed by Held and Shell (2012), and decompose the feedbacks into the Planck, lapse rate, relative humidity, and cloud components. The cloud feedbacks are computed by adjusting the cloud radiative effect for non-cloud influences (Shell et al., 2008; Soden et al., 2008), using the radiative kernels of Huang et al. (2017). The radiative kernels of Huang et al. (2017) are used because more models passed the clear-sky linearity test using those kernels (Zelinka et al., 2020).

3. Baseline Climate

A 1-year simulation is too short to evaluate all aspects of the climate, but an entire seasonal cycle provides a glimpse into the mean-state climate produced in a model. This section reviews aspects of the climate that are relevant to the radiative, hydrologic, and circulation response to warming to determine how the baseline simulation agrees with observations.

In a series of 40-day simulations, Donahue et al. (2024) reported that the global-mean, net top of atmosphere (TOA) radiative fluxes were within 1.2 W m^{-2} of satellite estimates (CERES-EBAF) across the four seasons. Since then, a few changes were made to the model to improve the physics and biases identified by Donahue et al. (2024) (see Section 2.1). After those updates, SCREAM 3 km produces a net global mean top of atmosphere radiative flux imbalance of 4.3 W m^{-2} , which is 3 W m^{-2} greater than CERES-EBAF Edition 4.2 (Table 2). For reference, the range of top of atmosphere radiative flux imbalance ranges from 0.0 to 3.5 W m^{-2} in CMIP6 models over the 1980–2008 period (Tables S1 and S5 in Supporting Information S1). The SCREAM 12 km initial condition ensemble shows that the uncertainty from internal atmospheric variability is around 0.3 W m^{-2} . While both shortwave and longwave flux biases contribute to the discrepancy between SCREAM 3 km and the satellite estimates, most of the difference comes from biases in the outgoing longwave radiation (OLR), which is too low in SCREAM 3 km. A further breakdown of the longwave flux into the clear-sky and cloud radiative effects shows that the difference is mainly due to the clear-sky fluxes, the cause of which is discussed later in this section.

In Table 2, we also see that SCREAM 12 km has a robustly more negative net TOA radiative flux compared to SCREAM 3 km. While there is a roughly 1.5 W m^{-2} larger net OLR in SCREAM 12 km than SCREAM 3 km, there is an even greater difference in the shortwave flux, where SCREAM 12 km reflects 4.7 W m^{-2} more shortwave to space, mainly due to cloud reflectance. This decreased reflection of clouds at higher resolution (SCREAM 3 km), despite no changes in the physics parameters, is consistent with the results of Hohenegger et al. (2020), who noted a weakening of shortwave cloud radiative effects with increasing resolution in the ICOSahedral Non-hydrostatic (ICON) model. We delve into the resolution sensitivity of clouds in Section 5.

Table 2

Global-Mean Radiative Fluxes at the Top of Atmosphere (TOA) and Precipitation Rates at the Surface for the Simulation Period of September 2019 to August 2020 in Observations, SCREAM 3 km, and SCREAM 12 km

Field	Obs	3 km	12 km	12 km ¹	12 km ²	12 km ³	12 km ⁴
Net TOA (W m^{-2})	1.3	4.3	−2.0	−1.8	−1.7	−1.7	−1.8
Net CRE	−19.2	−19.2	−20.2	−20.2	−20.2	−20.2	−20.2
Net clr-sky	20.7	23.4	21.7	21.7	21.8	21.9	21.8
SW TOA (W m^{-2})	242.4	243.1	238.2	238.4	238.5	238.5	238.5
SW CRE	−44.3	−44.2	−48.4	−48.2	−48.2	−48.2	−48.1
clr-sky SW	286.8	287.2	286.6	286.6	286.7	286.7	286.7
LW TOA (W m^{-2})	−241.1	−238.8	−240.3	−240.2	−240.2	−240.2	−240.4
LW CRE	25.1	25.0	24.6	24.7	24.7	24.7	24.6
clr-sky LW	−266.1	−263.8	−264.9	−264.9	−264.9	−264.8	−264.9
Precip (mm d^{-1})	2.88	3.03	3.15	3.14	3.14	3.14	3.15

Note. The four additional SCREAM 12 km ensemble members are distinguished by superscripts. For observational estimates, CERES-EBAF Ed 4.2 is used for radiative fluxes, and GPM-IMERG is used for precipitation rate. The sign convention of all the fluxes are such that downward fluxes are positive. The net TOA fluxes are broken down into shortwave (SW) and longwave (LW) components and into the clear-sky (clr-sky) and cloud radiative effect (CRE) contributions.

The global-mean radiative fluxes in Table 2 show that SCREAM 3 km does an adequate job of capturing the most basic statistic of Earth's radiative balance and cloud radiative effects (CRE). The maps of the TOA shortwave (SW) and longwave (LW) CRE in SCREAM 3 km and their differences (biases) compared to CERES-EBAF (Figure 1) show that despite relatively good agreement in the global mean and skill in representing most of the expected cloud patterns, some regional biases exist. For example, SCREAM 3 km tends to underestimate the SW CRE over many land regions, particularly over the northern hemisphere, and produces clouds which are more reflective over the Southern Ocean. And while SCREAM 3 km captures the LW CRE of high-clouds over most of the major convective regions, it tends to miss clouds over the northern part of the Indian Ocean and Tropical West Pacific, which also shows up in the SW CRE biases. On the other hand, SCREAM 3 km captures the subtropical marine stratocumulus relatively well: the strength of the SW CRE over the southeast Pacific and southeast Atlantic stratocumulus decks are similar to those of CERES-EBAF, though with a northward shift, while the SW CRE of the northeast Pacific stratocumulus are slightly too weak. The adequate representation of subtropical

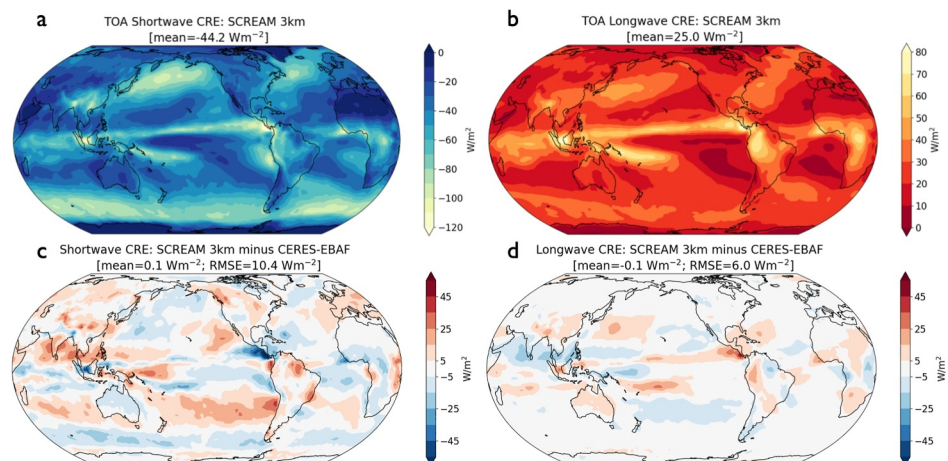


Figure 1. Present-day top-of-atmosphere, cloud radiative effect (CRE) in SCREAM 3 km and its difference with CERES-EBAF satellite estimates. (a) The baseline shortwave CRE in SCREAM 3 km. (b) The baseline longwave CRE in SCREAM 3 km. (c) The difference between CERES-EBAF and SCREAM 3 km in shortwave CRE. (d) The difference between CERES-EBAF and SCREAM 3 km in longwave CRE. To facilitate comparison, both SCREAM and CERES-EBAF retrievals are regridded to a uniform $2^\circ \times 2^\circ$ grid.

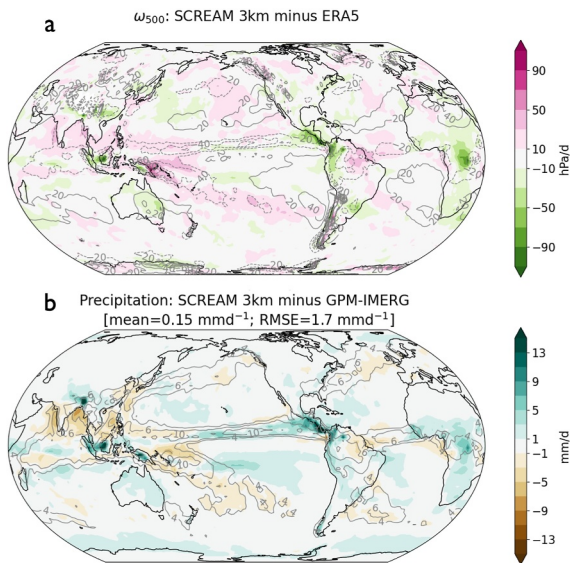


Figure 2. (a) Present-day difference in 500 hPa vertical pressure velocity (ω_{500}) between ERA5 and SCREAM 3 km. (b) Present-day difference in surface precipitation between GPM-IMERG and SCREAM 3 km. For reference, contours indicate baseline present-day values of ω_{500} from ERA5 and precipitation from GPM-IMERG. SCREAM, reanalysis, and satellite data are regridded to a coarser $2^\circ \times 2^\circ$ grid in these figures.

consistent with both vertical velocity and CRE differences (Figure 2b). The spatial RMSE of SCREAM 3 km's precipitation rate compared to GPM-IMERG retrievals for the same time period is 1.7 mm d^{-1} , which places it above the CMIP6 range 1.02 and 1.6 mm d^{-1} , where the CMIP6 RMSE values were computed from comparing each model's annual-mean precipitation rate against corresponding GPCP v3.2 retrievals between the years 2001–2008. As with the CRE comparisons, RMSEs are calculated after regridding both model and satellite data to a uniform $2^\circ \times 2^\circ$ grid.

One might speculate that the higher SW absorption over 30°S along with the higher reflectance around 50°S in Figure 1 might be due to a difference in the jet location. Indeed, the jet in SCREAM 3 km appears to be stronger and shifted poleward, and the ensemble of SCREAM 12 km models shows that the poleward displacement of the southern hemisphere midlatitude jet appears to be a shared feature across resolutions (Figure 3). Although consistent with the CRE biases, closer analysis is necessary to determine the contribution of the jet location in SCREAM 3 km to the shortwave cloud biases over the Southern Ocean. Figure 3 also shows that while SCREAM 3 km captures the shape of the upper level jet, the easterly flow over the equator is too strong compared to ERA5 estimates and outside of the envelope of what is seen in CMIP6 models.

Finally, we examine the temperature and humidity biases in SCREAM 3 km, with respect to ERA5 estimates (Figure 4). Many of the features, such as a generally warmer atmosphere, particularly in the Tropical lower and upper troposphere and the dry bias in the Tropical lower free troposphere, are consistent with what was reported by Donahue et al. (2024). The humidity biases in particular suggest an issue with SCREAM's convection, and are some of the largest persistent biases in the model. The relative humidity (RH) bias generally shows the same structure as the water vapor mixing ratio bias but also highlights moist biases in the upper troposphere. One feature that differs from those reported by Donahue et al. (2024) is the cold temperature bias around 200 hPa over the extratropics. The high RH and low temperature biases in the upper troposphere likely explain the clear-sky outgoing longwave bias reported in Table 2. Although a closer analysis is the subject of a follow-on study, the cold temperature is consistent with a lack of wave breaking at this location (Garcia et al., 2017) and might also explain the poleward bias of the midlatitude jets.

While a more detailed assessment of the cloud properties will be covered by Chao et al. (submitted to *JAMES*), overall, we find that SCREAM produces a present-day climate that is as skillful as CMIP models, with its specific areas of strength and areas needing improvement, particularly in the humidity distribution over the Tropics and

marine stratocumulus in SCREAM 3 km is also found in the km-scale simulations of ICON (Nowak et al., 2025; Segura et al., 2025), though a wide range of representation of stratocumulus exists among GSRMs participating in DYAMOND, implying that their representation are still highly dependent on parameterized processes (Heim et al., 2021). On the other hand, there is a band of positive SW CRE bias in SCREAM 3 km that wraps across the southern hemisphere oceans (Figure 1c). The spatial root-mean square error (RMSE) in SCREAM 3 km's SW CRE and LW CRE, when compared to CERES-EBAF, are 10.4 and 6.0 W m^{-2} . If we compare these values with RMSE values computed from comparing each CMIP6 model's (Table S1 in Supporting Information S1) annual-mean SW and LW CRE against CERES-EBAF's in each year between 2001 and 2008, SCREAM 3 km's SW CRE RMSE lies on the lower end of the CMIP6 range of 9.1 and 20.6 W m^{-2} , and the LW CRE RMSE lies below the CMIP6 range of 6.3 and 10.4 W m^{-2} . All RMSEs are calculated after regridding data to a uniform $2^\circ \times 2^\circ$ grid. This suggests that despite minimal tuning, the spatial distribution of clouds in SCREAM 3 km are better represented than most CMIP6 models.

Next, we examine the vertical velocity and precipitation differences in SCREAM compared to ERA5 and IMERG data (Figure 2). Consistent with the LW CRE biases, the large-scale ascent is overly weak over the Tropical West Pacific and northern Indian Ocean. On the other hand, SCREAM exhibits overly strong upward motion over central Africa, the northern extent of the Andes, the equatorial eastern Pacific, and west of the Panamanian gap (Figure 2a). The pattern and magnitude of precipitation differences are

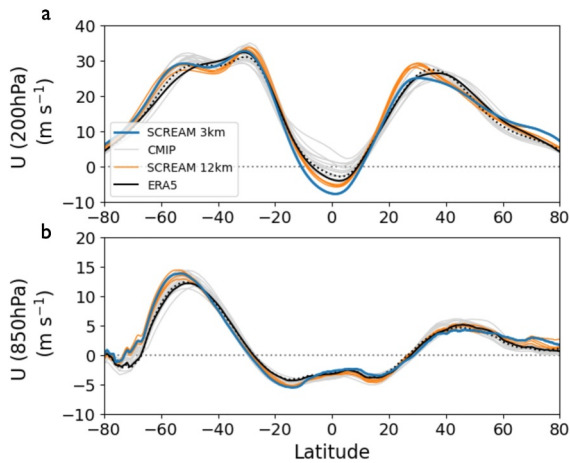


Figure 3. (a) Time- and zonally averaged zonal winds at 200 hPa in SCREAM 3 km (blue), SCREAM 12 km (orange), CMIP6 models (gray lines), and ERA5 (black line). The dotted and solid black lines indicate the zonal-mean from years 1979–2019 and September 2019 to August 2020, respectively. (bottom) Same as top plot, but for zonal winds at 850 hPa.

the spatial distribution of precipitation. In Section 8, we provide some discussion on potential reasons behind the mean-state biases and how some of the biases might impact the warming response described below.

4. Radiative Response to Warming

4.1. Global Averages

In the previous section, we examined how well SCREAM 3 km captures the present-day, mean climate. In the following sections we characterize large-scale climate response to a 4K warming in sea-surface temperatures, first focusing on the radiative response. Following the work of Cess et al. (1989) and subsequent studies (e.g., Qin et al., 2022; Ringer et al., 2014), we use a pair of experiments with prescribed, present-day sea-surface temperature (SST) and with the SST increased by 4 K to estimate the radiative feedbacks.

Figure 5 summarizes the global mean total radiative feedback and its breakdown into various components (the Planck, lapse rate, RH, albedo, and cloud feedbacks) and shows SCREAM 3 km's radiative feedback alongside estimates from CMIP5 and CMIP6 models. SCREAM 3 km's total radiative feedback lies in the range of CMIP models and within the top 15th percentile of radiative feedbacks, which suggests a larger than average equilibrium climate sensitivity (ECS) compared to CMIP class models. Also plotted

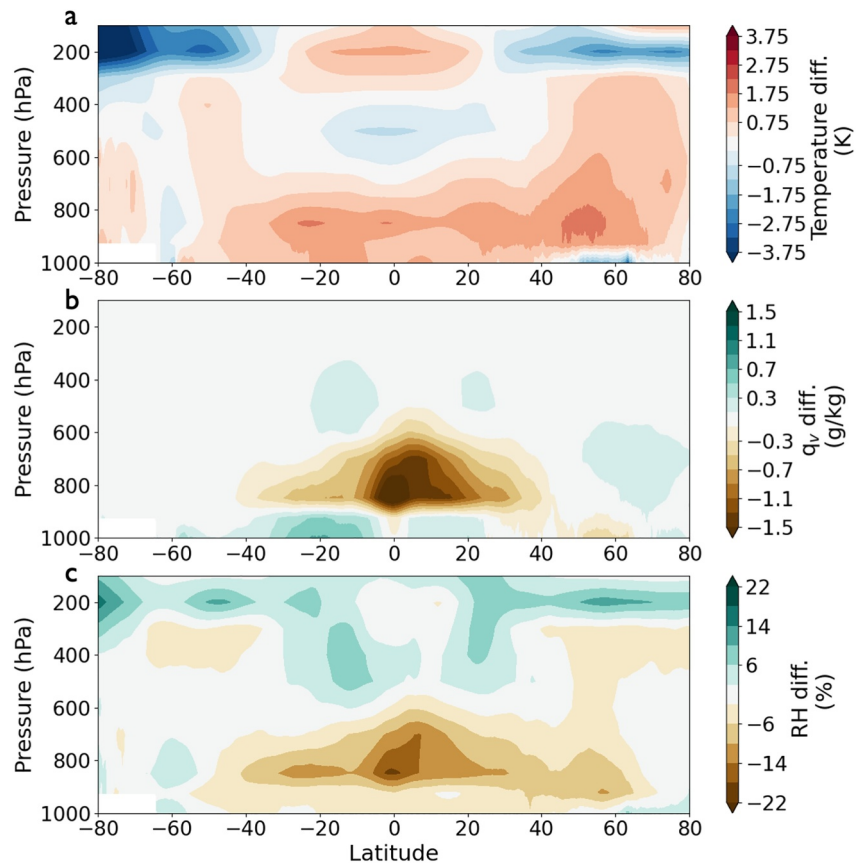


Figure 4. (a) Time- and zonally averaged difference in the temperature between SCREAM 3 km and ERA5 for the simulation period (September 2019 to August 2020). (b) Same as (a), but for the water vapor mixing ratio (q_v). (c) Same as (a), but for the relative humidity (%), derived from monthly temperature and water vapor mixing ratio values in both SCREAM 3 km and ERA5.

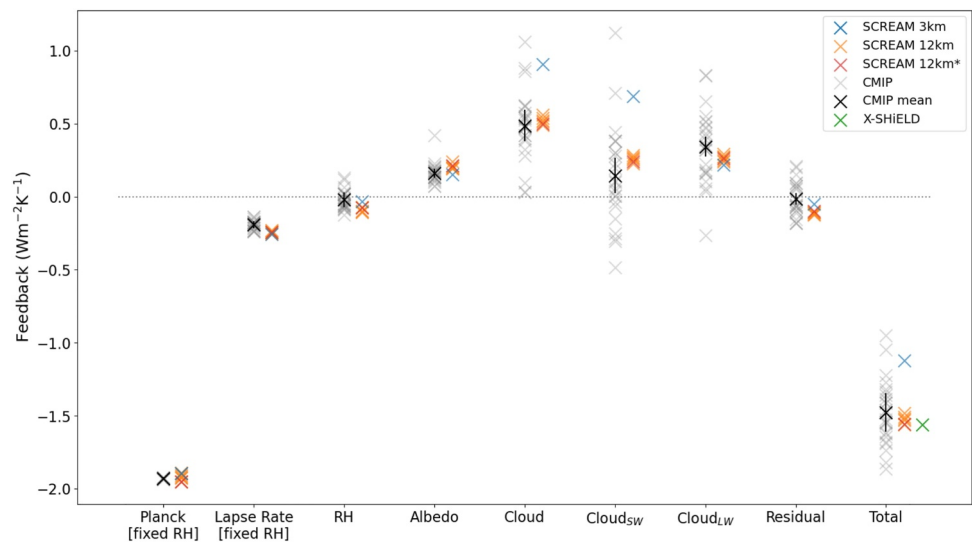


Figure 5. The global-mean radiative feedbacks in SCREAM 3 km (blue), SCREAM 12 km (orange), and CMIP5/6 models (gray). Also shown is the SCREAM 12 km simulation for a different time period (March 2020–February 2021) in red, the total radiative feedback from X-SHIELD as reported by Merlis et al. (2024) in green, and the CMIP5/6 ensemble mean in black. CMIP5/6 estimates are those calculated by Qin et al. (2022). The vertical line on the CMIP ensemble mean represents the interannual spread in the global mean feedbacks (see text for details). The Total feedback is categorized into the Planck, lapse rate (LR), relative humidity (RH), surface albedo, and cloud feedback terms, with the residual also indicated. The cloud feedbacks are further decomposed into the shortwave ($Cloud_{sw}$) and longwave ($Cloud_{LW}$) components.

alongside the SCREAM 3 km feedbacks (blue) are feedbacks from five ensemble simulations using SCREAM 12 km (orange) where the initial conditions have been slightly perturbed to provide a measure of the uncertainty from internal variability. The relatively tight spread ($0.2 \text{ W m}^{-2} \text{ K}^{-1}$) among the five simulations suggests that our constraint of only being able to run one 13-month simulation should have a small impact on the uncertainty of the feedback estimates. Furthermore, calculating the Cess feedbacks over a different period with SCREAM 12 km (February 2020–March 2021—in red), does not appear to change the feedback appreciably. Although we have not performed a multi-year simulation with SCREAM 3 km, as is typically done in CMIP's AMIP protocol, the spread in the interannual difference in radiative feedbacks, represented by the vertical bar on the black X, indicates that spread coming from year-to-year variability is also small compared to the intermodel spread in radiative feedbacks (Qin et al., 2022). The vertical line on the CMIP ensemble mean represents the inter-model median range of the interannual spread in the global mean feedbacks. In other words, the global-mean feedback is calculated from each CMIP model's amip, amip+4K pair for each simulated year to calculate the range (difference between the highest and lowest) in the annually calculated feedback. Then the median value from across the models is shown as the black line. Despite our attempts to provide a best estimation of the uncertainties due to the short 1-year simulation, we still acknowledge that the effect of internal variability and period dependence of the feedbacks in SCREAM 3 km might differ from the SCREAM 12 km or CMIP models and so the true uncertainties from running just 1 year with SCREAM 3 km are unknown until we perform a multi-year Cess experiment with SCREAM 3 km.

A relatively large ($0.45 \text{ W m}^{-2} \text{ K}^{-1}$) difference in the total feedback exists between SCREAM 3 km and X-SHIELD, another km-scale GSRM without any deep convective scheme, run for an overlapping period (Merlis et al., 2024). The difference in the total feedback between SCREAM 3 km and SCREAM 12 km is also of similar size ($0.4 \text{ W m}^{-2} \text{ K}^{-1}$), which indicates a substantial resolution sensitivity in SCREAM's radiative feedbacks, despite no changes in the physical formulation or tuning parameters. Other studies have noted specific feedbacks that show resolution sensitivity (Blossey et al., 2009), and here we show that spread can be quite large in the total feedback estimate.

We next look at how individual feedbacks contribute to the total feedbacks and their intermodel differences. For the breakdown of the total feedback, we use the constant relative humidity framework proposed by Held and Shell (2012). In this framework, the Planck and lapse rate feedbacks include the radiative impact of water vapor

changes necessary to maintain constant relative humidity as the temperature changes, following Clausius-Clapeyron. As Held and Shell (2012) note, such a framework removes the canceling impacts of the conventional lapse rate and water vapor feedbacks and highlights where the true uncertainties exist in multi-model analysis.

The negative Planck feedback is the largest in magnitude, whereas the negative lapse rate feedback is of similar magnitude to the positive albedo feedback. The magnitude of the relative humidity feedback is small, although the CMIP intermodel spread of $0.25 \text{ W m}^{-2} \text{ K}^{-1}$ stretches into both positive and negative values. Unsurprisingly, the largest intermodel spread is found in the cloud feedback, with a range exceeding $1.0 \text{ W m}^{-2} \text{ K}^{-1}$. A larger contribution to the uncertainty stems from the shortwave cloud feedback, though the spread in longwave cloud feedback is also considerable.

SCREAM 3 km differs most from the CMIP multi-model mean in its cloud feedback, with a total cloud feedback strength of $0.9 \text{ W m}^{-2} \text{ K}^{-1}$. The decomposition into shortwave and longwave cloud feedbacks reveals that SCREAM 3 km's positive shortwave cloud feedback is one of the strongest compared to CMIP models, while the positive longwave cloud feedback is slightly below the CMIP multi-model mean. Compared to the CMIP ensemble, SCREAM 3 km has a weaker negative Planck feedback than all CMIP models, one of the strongest negative lapse rate feedbacks, and a near zero RH feedback, which is close to the CMIP multi-model mean. In these atmosphere-only warming experiments where the sea-ice extent is held fixed, the surface albedo feedback is also close between SCREAM 3 km and the CMIP mean.

Compared to SCREAM 3 km, SCREAM 12 km produces similar feedbacks across most of the feedback components. The 3 and 12 km differ in the shortwave cloud feedback, where SCREAM 12 km produces a feedback much closer to the CMIP ensemble mean. There are slight ($<0.06 \text{ W m}^{-2} \text{ K}^{-1}$) differences between SCREAM 3 km and SCREAM 12 km for other feedbacks; Planck and RH feedback are more negative in SCREAM 12 km (though the spread in Planck feedback among SCREAM 12 km experiments indicates differences between SCREAM 3 km, SCREAM 12 km, and CMIP models might not be robust), while the albedo feedback is more positive in SCREAM 12 km, but the differences do not contribute to the large difference in the total feedback. As is the case with the total feedback, the spread among the five SCREAM 12 km simulations is relatively small across the various breakdowns in the feedbacks, which supports the robustness of the global-mean differences that we see between the SCREAM 3 km and SCREAM 12 km simulations.

4.2. Spatial Structure

Because the SCREAM 3 km shows such a large difference in the global mean cloud feedbacks when compared to the CMIP multi-model mean and SCREAM 12 km, we next examine the spatial characteristics in the cloud feedbacks. Despite differences in the global mean values, Figure 6 shows that there are notable similarities in the spatial structure of the cloud feedbacks across SCREAM 3 km, SCREAM 12 km, and the CMIP multi-model mean. First, there are positive shortwave cloud feedbacks found in the subtropical marine stratocumulus regions across Figures 6a–6c and in the northern and southern hemisphere storm tracks. Both SCREAM and the CMIP multi-model mean also show strongly positive shortwave feedbacks over the southern half of the African continent. There are also regions of negative shortwave feedbacks, namely over the Southern Ocean, roughly centered around 60°S , and over the Tropical Pacific, spanning southeastward from just north of Maritime Continent toward the South Pacific Convergence Zone.

In the longwave, all models show a positive feedback over the Tropical West Pacific, around the Maritime Continent, although the models differ in the exact locations. To what extent the difference in the patterns of positive and negative longwave cloud feedbacks in the Tropical Pacific are due to internal variability or model differences remains to be seen with longer simulations of SCREAM. Longwave cloud feedback is positive in most regions, with distinct areas of negative feedback, namely over the African continent and over the Amazon that coincide with positive shortwave feedbacks.

We might have expected differences between SCREAM 3 km and SCREAM 12 km (or the CMIP-mean) to mainly appear in the deep convective regions in both the longwave and shortwave cloud feedback, given the impact of resolution on the nature of resolved deep convection, but we do not observe as strong a resolution sensitivity in the longwave feedback, compared to the shortwave feedback, which suggests that the differences stem from low-level clouds. The magnitudes of positive feedback in SCREAM 3 and 12 km over the subtropical

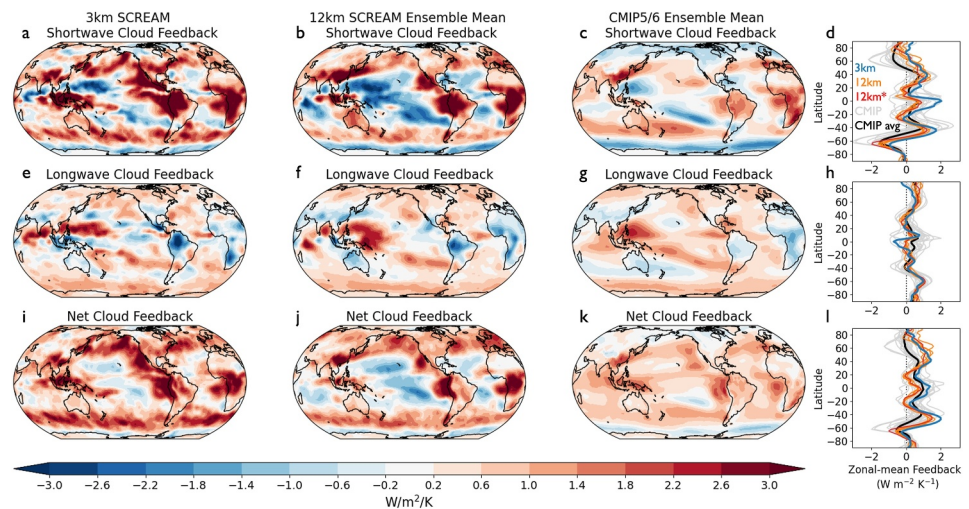


Figure 6. The spatial map of the shortwave and longwave components and their combined net cloud feedback (Cloud_{sw} , Cloud_{lw} , Cloud of Figure 5) in SCREAM 3 km (a, e, i), ensemble of SCREAM 12 km (b, f, j), and CMIP5/6 ensemble mean (c, g, k). Also shown on the far right column (d, h, and l) are the zonal means of the cloud feedbacks from SCREAM 3 km (blue), SCREAM 12 km (orange), CMIP models (gray), and the CMIP ensemble mean (black). The zonal mean of the SCREAM 12 km simulation (SCREAM 12 km*) covering a different time period (March 2020–February 2021) is also shown in red.

marine stratocumulus are larger than the CMIP multi-model mean. SCREAM 3 km's negative shortwave feedback over the Southern Ocean is also one of the weakest, most evident the zonal mean plots in Figure 6d. Furthermore, over the northern hemisphere mid- and high-latitudes, the zonal mean shortwave cloud feedback is positive in SCREAM 3 km, whereas most CMIP models show a negative feedback. While there are some resolution sensitivities between SCREAM 3 km and SCREAM 12 km over the mid- and high latitude regions, with SCREAM 3 km exhibiting a more positive feedback compared to SCREAM 12 km, the largest differences are over the tropical and subtropical oceans. Consistently, SCREAM 12 km exhibits a more negative cloud feedback over the west and central tropical and subtropical ocean basins and a weaker positive feedback over marine stratocumulus regions compared to SCREAM 3 km (Figures 6i and 6j). The cloud differences that lead to the differences in the radiative feedback are discussed in Section 5 and their correspondence with previous published studies of km-scale global models is discussed in Section 8.

5. Process Understanding of Low-Cloud Response to Warming

Previous analysis has revealed substantial differences in the SW cloud feedback between the 12 and 3 km simulations in SCREAM. To further investigate these differences, we delve deeper into the behavior of low clouds, particularly focusing on their response in the tropical and subtropical regions within these simulations. Figure 7 presents the cloud fraction profiles for both stratocumulus and cumulus regimes across the two resolutions. The distinction between these regimes is based on conditional sampling following the methodology of Medeiros and Stevens (2011), with further details provided in the caption of Figure 7.

The results indicate a clear resolution sensitivity in how these low clouds are represented in the control configurations, with differences observed in both cloud amount and cloud thickness. Specifically, the lower resolution 12 km simulation tends to produce thicker clouds with a higher cloud fraction compared to the 3 km simulation. This resolution sensitivity aligns with previous findings (Bretherton & Khairoutdinov, 2015; Hohenegger et al., 2020), which suggest that global cloud-resolving models (GCRMs) generally generate more low clouds at lower resolutions compared to their operational counterparts. One possible explanation, proposed by Blossey et al. (2009), is that higher resolution better resolves shallow convective updrafts, leading to stronger entrainment and boundary-layer ventilation that dries and thins the cloud layer, thereby reducing cloud fraction and cloud depth. This mechanism could contribute to the differences seen between the 12 and 3 km SCREAM

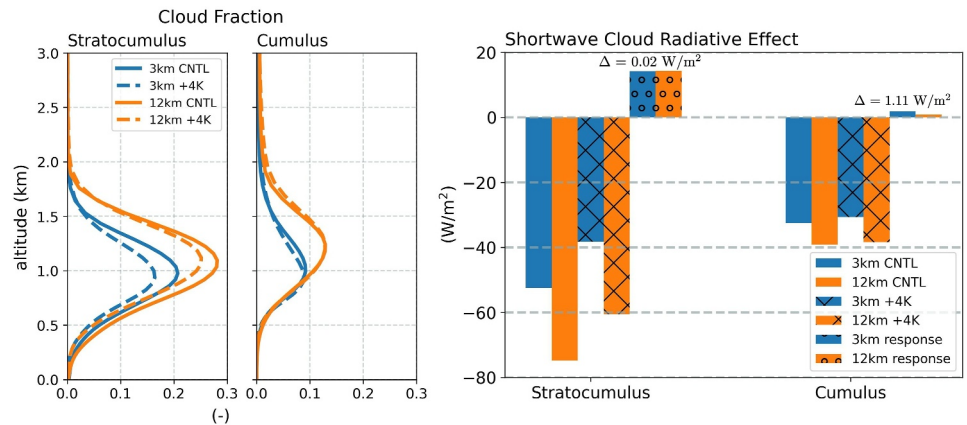


Figure 7. Left two panels show averaged conditionally sampled profiles of cloud fraction over the subtropical ocean while the right panel represents the conditionally averaged shortwave cloud radiative effects for stratocumulus and cumulus. The conditional sampling between stratocumulus and cumulus is based on Medeiros and Stevens (2011). For both regimes the latitude is bound by 30°S and 30°N , only oceanic columns are considered (ocean fraction equal to 1), and $\omega_{500} > 10 \text{ Pa s}^{-1}$. For Stratocumulus the lower tropospheric stability (LTS) is required to be $> 18.5 \text{ K}$ while for Cumulus $18.5 > \text{LTS} > 11.3 \text{ K}$.

simulations, although observations from a recent field campaign over the trade wind region point to the importance of mesoscale vertical motions in explaining variations in the cloud fraction of trade wind cumuli (Vogel et al., 2022).

This resolution sensitivity is somewhat unexpected for SCREAM. Previous studies by Bogenschutz et al. (2023) demonstrated that SCREAM exhibited robust behavior to resolution changes for subtropical and Arctic stratocumulus clouds in idealized, short-duration simulations. However, more recent work by Bogenschutz et al. (2025) has shown that this resolution sensitivity becomes more pronounced when time-varying forcing (in large-scale advective tendencies and surface conditions) is introduced in longer-duration simulations, consistent with the findings of the present study.

The pronounced resolution sensitivity in low clouds within SCREAM also implies that the differing responses to increased sea surface temperatures (SSTs) are not surprising. The cloud fraction and SWCRE response panels in Figure 7 illustrate that stratocumulus clouds in SCREAM are generally more responsive to warming than cumulus clouds. However, the responses of stratocumulus clouds between the 12 and 3 km resolutions are nearly identical, with only a 0.02 W m^{-2} difference between the simulations. This small difference is likely due to the cancellation of regional variations, as some stratocumulus regions, like California, exhibit a strong positive response difference, while others, such as the Namibian region, show a negative response of nearly equal magnitude (Figure 6).

In contrast, the cumulus regime, while less responsive to SST changes than stratocumulus, displays a larger response difference between resolutions, with a 1.11 W m^{-2} difference. Since cumulus clouds cover a much larger area of the subtropical oceans (45% vs. 4% in SCREAM simulations), their response plays a crucial role in driving the overall response differences between the 12 and 3 km resolutions. The cloud fraction profiles further suggest that, with warming, the cumulus clouds in the 12 km simulation become slightly deeper, while those in the 3 km simulation become shallower. This behavior appears reproducible in doubly periodic SCREAM (DP-SCREAM) simulations, suggesting that the resolution dependence of cumulus cloud depth is more strongly tied to local cloud-scale dynamics and parameterization processes than to differences in large-scale subsidence. Further investigation is needed to fully diagnose the physical mechanisms involved. Bogenschutz et al. (2023, 2025) note that shallow cumulus clouds are particularly sensitive to resolution in SCREAM and are not properly represented, unless the model is run with LES-like resolutions, often being too shallow and resembling stratocumulus clouds. This is evident in Figure 7, where the vertical extent of cumulus clouds in SCREAM is nearly identical to that of stratocumulus clouds. These findings underscore the need for improvements in representing shallow convection in SCREAM, potentially through coupling the SHOC parameterization with a mass flux approach, as suggested by Chinaita et al. (2023).

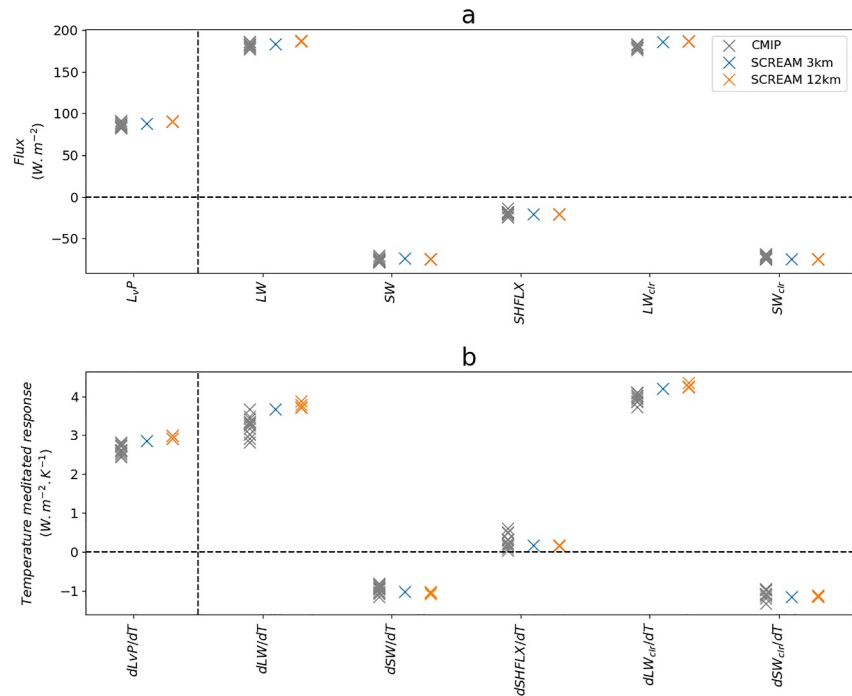


Figure 8. (a) A breakdown of the atmospheric energy budget in the baseline simulation coming from latent heating (L_vP), longwave cooling (LW), shortwave absorption (SW), and sensible heat flux (SHFLX). The clear-sky contribution to LW and SW (LW_{clr} and SW_{clr}) are also shown. Convention is positive for radiative cooling, negative for radiative warming, and negative for upward SHFLX into the atmosphere. CMIP6 models are indicated in gray, SCREAM 3 km in blue, and SCREAM 12 km in orange. (b) The temperature-mediated response of the same atmospheric energy budget terms.

Finally, the resolution sensitivity observed in the stratocumulus regime within SCREAM warrants further examination. This sensitivity is reproducible in DP-SCREAM, which could accelerate understanding and lead to parameterization improvements, as suggested by Bogenschutz et al. (2025).

6. Large-Scale Hydrological and Circulation Response to Warming

Our year-long integrations allow us to quantify our GSRM's hydrological response to warming. Specifically, we can conduct an energetic breakdown to quantify how mean precipitation responds to the 4 K increase in sea surface temperature as in DeAngelis et al. (2015). Before we delve into the temperature-mediated response of mean precipitation, it is worthwhile to examine the atmospheric energy budget in the baseline state. This may shed light on what drives SCREAM's precipitation bias (Table 2) and inform an assessment of the temperature-mediated response. To that end, Figure 8a shows mean precipitation (in units of W m^{-2}) and its energetic contributions in CMIP6 models, SCREAM 3 km, and the SCREAM 12 km ensemble. Overall, SCREAM falls within the CMIP6 range for most energy fluxes. The exception is clear sky longwave cooling, where all SCREAM configurations tend to lose radiation in the clear sky at a rate slightly higher than CMIP. We believe this is fundamentally tied to the temperature and moisture biases shown earlier. At the top of the troposphere, positive moisture and temperature biases can compensate and lead to a TOA flux that is within the CMIP range. On the other hand, the positive temperature bias in the lower troposphere increase the loss of energy to the surface. The effect of this bias is thus atmospheric radiative cooling that is on the higher end of the CMIP range. However, the all sky longwave radiative cooling is within the CMIP6 range, indicating that clouds compensate for some of this excess loss.

All SCREAM simulations exhibit a higher temperature-mediated response of precipitation compared to CMIP (Figure 8b). We can see that this is driven by an anomalously high longwave cooling response which in turn is driven by an anomalously high clear sky longwave cooling response. It is possible that the high longwave cooling response stems from a high longwave cooling rate in the baseline climate. Most notably, longwave emission by clouds in the baseline climate, which acted to compensate for the excess loss in the clear sky, responds more

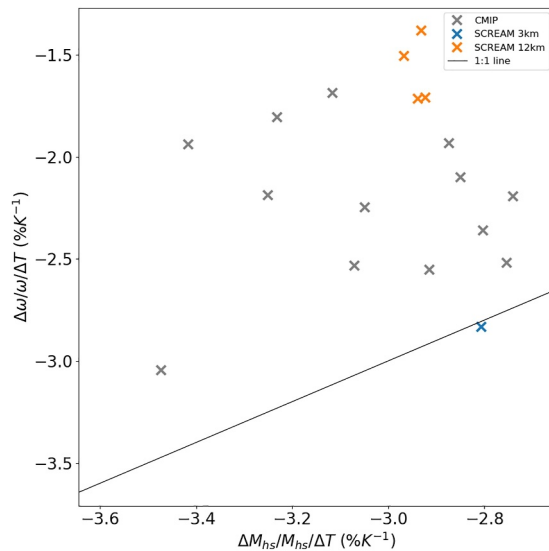


Figure 9. Relative change in area weighted mean upward velocity at 500 hPa against the Held and Soden (2006) proxy for convective mass flux. Gray x's represent CMIP6 models while the blue and orange represent SCREAM 3 km and the SCREAM 12 km ensemble respectively. The tropical band of 30°S to 30°N is used for this analysis.

weakly than its clear sky counterpart. This translates to a high precipitation response even though the baseline precipitation is within the CMIP range. We therefore suspect that SCREAM's strong temperature mediated-precipitation response is rooted in our model's tendency to have a warmer boundary layer though a further investigation is required to confirm this. Indeed, we see this feature in both SCREAM 3 and 12 km, which indicates that resolution plays a small role in dictating the hydrologic response to warming.

Next, we examine whether SCREAM, which relies on the explicit simulation of deep convection, predicts a tropical circulation response that is similar to CMIP. We also examine whether the tropical circulation response is related to the hydrological sensitivity as conjectured by Vecchi and Soden (2007). To that end, we plot the relative change in tropical area-weighted mean upward pressure velocity at 500 hPa against a proxy for the convective mass flux in Figure 9. The upward pressure velocity is computed by averaging over all grid cells where the monthly mean pressure velocity at 500 hPa is positive and normalizing by area. We use the Held and Soden (2006) proxy for convective mass flux which is the ratio of mean tropical precipitation to boundary layer moisture. Some notable features of the scatter plot are the lack of a predictive correlation between the two metrics and the fact that convective mass flux systematically overpredicts the circulation response. This could be attributed to the tightness of the spread in both metrics—a consequence of the fact that these models are using identical SST forcings. While Vecchi and Soden (2007) found relative success in relating the mean ascent to the

convective mass flux proxy, it is possible that the utility of the approach is lost when using same SST forcings. In other words, the predictive capability is lost because the spread becomes smaller than the errors introduced in making the crude assumption of equating mean ascent upward velocity to a rough proxy of convective mass flux. Nevertheless, we note a model consensus on a 2%–3%/K weakening of the overturning circulation which can be broadly understood as a necessary consequence of the stronger increase in boundary layer moisture compared to precipitation, if we assume a tight relationship between SST and BL humidity. Looking more closely at SCREAM, we can see a substantial resolution sensitivity of the circulation response that is not mirrored in the mass flux proxy—further emphasizing the limits of a Held and Soden type analysis for understanding the circulation response in these idealized forcing experiments. The difference between the two resolutions in the circulation response is roughly equal to the CMIP range, though more years with SCREAM 3 km are necessary to determine the significance of this correspondence. SCREAM 3 km also sits almost perfectly on the 1:1 line but given the spread seen in SCREAM 12 km, it is also perhaps fortuitous.

Figure 9 also allows a comparison of the hydrological sensitivity among models. Since the SSTs are prescribed, the change in the convective mass flux proxy is dictated exclusively by the hydrologic sensitivity. We see that SCREAM does not exhibit anomalously high hydrologic sensitivity—in contrast to its temperature mediated response to precipitation. This is due to the fact that SCREAM's base state has a particularly high precipitation rate (as discussed previously), which then drives an anomalously high temperature mediated response of precipitation (with units of $\text{W m}^{-2} \text{K}^{-1}$), but which also places its percent change in precipitation rate with warming ($\% \text{K}^{-1}$ for hydrologic sensitivity) within the inter-quartile range of CMIP6 models.

The poleward shift of the midlatitude jet and Hadley cell that we see in CMIP models is also found in SCREAM and its shift is within the spread of CMIP6 models (Figure 10). Starting with the midlatitude jet, we note again SCREAM's tendency to have the baseline SH jet poleward compared to CMIP6 models for both 3 and 12 km simulations (also in Figure 3). Despite its baseline difference, its poleward shift is unremarkable and within the range of what is seen in CMIP6 models. The NH jet location, which is more difficult to identify than the SH jet in the annual mean (see Figure 3), has a larger poleward shift than the SH shift. We also note the large range of poleward shift in the SCREAM 12 km ensemble, which indicates the potential difficulty of using a 1-year simulation to accurately determine the forced response of the NH jet. We also see a poleward shift in the Hadley cell edge but note that the starting location and expansion in SCREAM 3 and 12 km are within the spread of

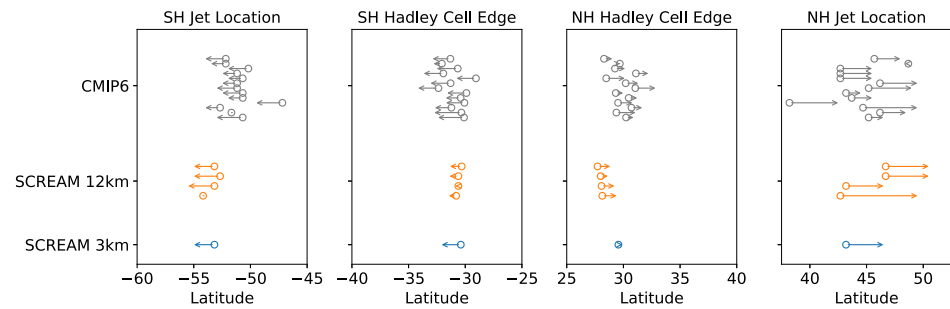


Figure 10. The expansion of the midlatitude jet and Hadley cell edge in the Northern and Southern Hemisphere (NH and SH) in CMIP6 (gray), SCREAM 3 km (blue), and SCREAM 12 km (orange) models. The midlatitude jet location is defined as the latitude of maximum zonally- and annually averaged zonal winds at 850 hPa in each hemisphere. The Hadley cell edge is determined using the mass stream function calculated from vertically integrating the meridional wind and identifying the latitude where the zero isopleth crosses 500 hPa in each hemisphere. Circles indicate the latitude in the baseline simulation and the arrows indicate the expansion due to the 4 K warming in SSTs.

CMIP6 models and no resolution sensitivity can be discerned. Overall, it appears that the circulation response in SCREAM largely agrees with what we see in CMIP6 models without a remarkable difference to the CMIP-mean response. We acknowledge that a fair amount of inter-annual spread in the jet latitude (roughly 8°) and Hadley cell latitude ($2\text{--}4^\circ$) exist (Kang & Polvani, 2011; Lucas et al., 2014), so interpreting temperature-mediated shifts should be done cautiously, and quantitatively narrowing the spread will likely be difficult with 1-year simulations.

7. Organized Convective Storms Response to Warming

Analysis of Cess-Potter experiments with GSRMs need not be limited to the large-scale atmospheric response to warming. Meso-scale convective systems (MCSs) play a non-negligible role in the global hydrological cycle and account for 24%–44% of global mean precipitation annually and for $\sim 50\%$ of 99th percentile extreme precipitation amounts, particularly in the tropics (Wu et al., 2024; Zhao, 2022). Present CMIP-class models with horizontal resolutions optimized for centennial scale experiments lack the resolutions to faithfully simulate MCS features. Models generally require horizontal resolutions of 4 km or finer to properly simulate most MCS features (Prein et al., 2017, 2021). Therefore, our Cess-Potter experiments with SCREAM 3 km provide a unique opportunity to investigate how MCS precipitation may respond to future warming.

We conducted MCS tracking using the PyFLEXTRKR (Feng et al., 2023) feature tracking algorithm to identify MCS events. We use the period from 1 December 2019 to 31 August 2020 for analysis in part because many MCS features occur over land, which necessitates a longer spin up to allow for soil moisture adjustment, but also because there are some gaps in the hourly output of the first 4 months.

Table 3 summarizes the average characteristics of total and MCS precipitation from the control and +4K simulations. We can see a general increase in the MCS precipitation amount ($\sim 22\%$) and in MCS precipitation fraction (with respect to total precipitation, 2.3%) globally with warming in the last 9-months of the SCREAM Cess-Potter experiment. The global mean MCS frequency is slightly decreased (0.29%).

Table 3

Average Characteristics of Total and MCS Precipitation From Cess Control and +4K SST Experiments

Mean values (60S–60N)	Cess control	Cess +4K
Total precipitation (mm day^{-1})	3.28	3.70
MCS precipitation (mm day^{-1})	1.47	1.79
MCS precipitation fraction (%)	31.64	33.94
MCS frequency (%)	4.18	3.89

Note. The period of analysis is from 00Z 1 December 2019 to 00Z 31 August 2020.

On the other hand, the regional response of MCS precipitation to the warming can be larger, with significant spatial variability (Figure 11). Compared to the control simulation, large areas of the globe show decreases in MCS precipitation frequency in the +4K simulation (Figure 11d), but with notable exceptions over portions of the Intertropical Convergence Zone and Southern Pacific Convergence Zone. Over land, MCS precipitation frequency generally decreases but MCS precipitation amount generally increases with warming (Figure 11e). For example, MCS frequency decreases over most of the Amazon and but precipitation amount is unchanged or increases. Over the continental US, MCS frequency also decreases, but precipitation amount is mostly unchanged. Regions with increased or unchanged MCS precipitation amount but decreased frequency imply fewer but stronger MCS events in response to warming. The spatial pattern of MCS fraction changes

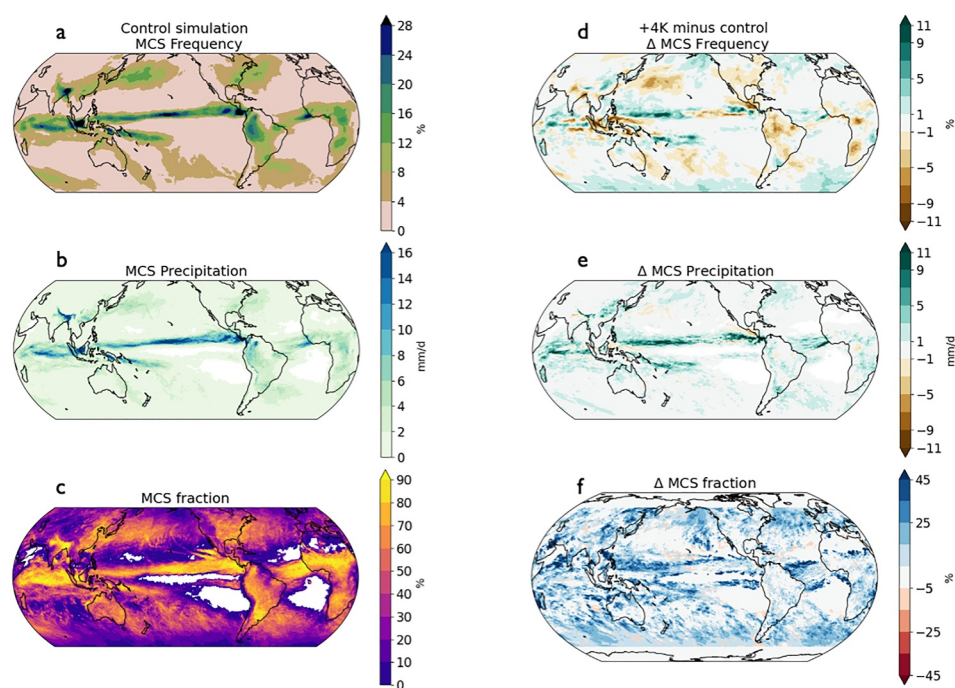


Figure 11. Mesoscale Convective System precipitation frequency (a, %), amount (b, mm day^{-1}), and fraction (c, %) from the Cess control simulation. Also show in (d–f) are the differences between +4K SST and control simulations for the same quantities. The period of analysis is from 00Z 1 December 2019 to 00Z 31 August 2020.

(Figure 11f), which show a widespread increased contribution of MCS to the total with warming, is generally similar to that of MCS amount changes (Figure 11e).

While these results are interesting, we acknowledge that longer simulations are needed to establish whether these MCS precipitation changes to SST forcing are robust in this model, and deeper analysis is needed to characterize the underlying mechanisms for each region. Furthermore, investigating MCS precipitation across several GSRMs would be useful for establishing whether this signal may be relevant to the real atmosphere, or whether the response is specific to this model. Nevertheless, this analysis demonstrates the value of GSRM experiments that can represent a response to warming over a range of scales from global to convective.

8. Discussion and Conclusions

This manuscript provides a broad overview of the large-scale features associated with the atmospheric radiation, water cycle, and dynamics and their response to warming in a pair of Cess-Potter simulations performed using the global km-scale atmosphere-model SCREAM. Despite performing very little tuning, SCREAM shows relatively good skill in representing the top of atmosphere (TOA) radiation. The simulated annual and global-mean TOA radiation fluxes are all within 3 W m^{-2} of satellite estimates (Table 2) and the overall spatial patterns of cloud radiative effect (CRE) are well captured by SCREAM, particularly over the subtropical eastern ocean basins where marine stratocumulus are typically found. The zonal winds in the lower and upper troposphere are also generally well captured, despite a poleward displacement of the Southern Hemisphere jet that might be due to a lack of upper atmosphere wave breaking in the Southern Hemisphere. On the other hand, we see some of the largest biases in shortwave and longwave CRE over the Tropics, particularly over land, the northern Indian Ocean, Tropical West Pacific Ocean, and Eastern Pacific Ocean. These regions also correspond to regions with the largest precipitation and vertical velocity biases. As with the humidity and temperature profile biases, we believe these stem from challenges in accurately simulating convection at these coarsely resolved (3.25 km) scales. Ongoing tests indicate that an improved treatment of rain fraction and its impacts on evaporation lead to notable improvements to the dry bias over the Tropical West Pacific, but it remains to be seen whether it will completely fix the biases or whether additional physics are needed, such as the representation of unresolved

entrainment and detrainment, as examined by Takasuka, Kodama, et al. (2024). Despite these biases, we find that SCREAM 3 km produces a realistic climate comparable in skill to 100 km-scale climate models, which have been extensively tuned to match the observed climate. SCREAM produces a better spatial RMSE in SW and LW CRE than most CMIP6 models examined, but produces a worse spatial RMSE in precipitation rate. In terms of possibility of future improvement, this is quite promising, given that these Cess-Potter experiments are the first time that SCREAM 3 km has been run out for a year.

Shifting our attention to the radiative and precipitation response of SCREAM to warming, we find that SCREAM 3 km's radiative feedbacks lie within the range of CMIP model estimates, although a strong shortwave cloud feedback puts its total feedback within the top 15% of CMIP5/6 models and much stronger than a previous estimate of total radiative feedback from the X-SHIELD model (Merlis et al., 2024), which was run at comparatively high resolution (3.25 km). Furthermore, SCREAM's precipitation response to warming is one of the highest among CMIP6 models and is due to a strong increase in longwave atmospheric cooling at the surface.

Five pairs of SCREAM 12 km Cess-Potter simulations provide some context for these results. First, they indicate that the impact of internal variability from weather noise is likely relatively small compared to the intermodel spread in most cases, though perhaps not for shifts in the Hadley cell and midlatitude jet. When we extend the SCREAM 12 km simulation out to a period encompassing March 2020 to February 2021, aligning with the year-long DYAMOND proposed by Takasuka, Satoh, et al. (2024), we find that choosing a different time period has minimal impact beyond the spread we see from weather noise. Although these results are not fully conclusive, because the 3-km configuration might exhibit a different internal variability and sensitivity to the period of simulation, they broadly support the findings of Qin et al. (2022) and Figure 5, indicating that the interannual spread in feedbacks is small compared to the inter-model spread across CMIP-class models. Similarly, Merlis et al. (2024) and Guendelman et al. (2024) also report using the X-SHIELD model that global mean estimates of radiative and precipitation response to warming are small when comparing 1 year to the next year. This gives us some confidence that the results from a pair of 1-year Cess-Potter experiments can tell us much about the radiative and hydrological response to warming in these idealized experiments. Second, the differences between SCREAM 3 km and SCREAM 12 km show a resolution sensitivity of the shortwave cloud feedback, and hence to the total radiative feedback. Analyses of simulations at intermediate resolutions are currently underway to better understand the resolution sensitivity of the various warming responses, and it is possible that even at 3 km, SCREAM's radiative feedback estimates continue to change with increasing resolution and higher resolutions are necessary to observe the convergence with resolution.

One might ask how the biases in SCREAM's mean-state impact the temperature response. As was discussed in Section 6, we find that the strong clear-sky downward longwave in SCREAM, which contributes to the higher global mean-state precipitation rate (Table 2), likely explains the large precipitation increase with warming. SCREAM also exhibits a large bias in the spatial distribution in precipitation, particularly a dry bias over the Tropical West Pacific. While we do not see a large difference between SCREAM and the CMIP multi-model mean in the net cloud feedback over the Tropical West Pacific, there are some differences in spatial patterns and the bias has the potential of impacting non-local cloud feedbacks over the tropical and subtropical oceans by impacting the mean-state and change in boundary layer-top inversion strength, which are crucial for subtropical cloud feedbacks (Qin et al., 2024; Qu et al., 2014; Sherwood et al., 2014). While sensitivity experiments to examine the impact of the mean-state biases on the climate response are difficult given the computational cost and beyond the scope of this study, preliminary experiments with a doubly periodic version of SCREAM suggest the resolution sensitivity in low-cloud feedbacks are due to local-scale physics, and future experiments with improved physics that are currently being prepared might shed light on the impact of improving some of these biases on the warming response found here.

Given the result where we find a strongly positive total feedback compared to those of Merlis et al. (2024), we next discuss our results in context of the existing literature on cloud feedbacks in GSRMs. One suggestive result coming from earlier studies is that km-scale simulations predict a more negative cloud feedback compared to conventional climate models. Miura et al. (2005), who ran NICAM with a dx of 7 km in an idealized aquaplanet-like configuration, reported an increase in cloud cover over the mid- and high-latitudes that leads to a globally averaged negative shortwave cloud feedback. Narenpitak et al. (2017) noted an increase in cloud cover over the

subtropics with warming in their aquaplanet-like configuration using SAM (dx of 4 km). In real geography simulations using the Super-Parameterized Community Atmosphere Model (SPCAM), Wyant et al. (2006) also reported a more negative cloud feedback than most climate models due to an increase in cloud fraction over the high-latitudes and an increase in the cloud fraction and condensate loading in low-level clouds in tropical subsidence regions. Analyses are underway to determine whether the mechanisms for the negative cloud feedbacks in SCREAM are the same as those proposed by Wyant et al. (2009). Similar to the resolution sensitivity we find in SCREAM, Blossey et al. (2009) found that the strength of the negative cloud feedbacks of Wyant et al. (2009) were reduced when cloud-resolving column modeling simulations with the SPCAM were run with finer (~ 100 m) resolution instead of the 4 km resolution used by Wyant et al. (2009). Although Merlis et al. (2024) do not report on individual feedbacks, the total feedback of $-1.56 \text{ W m}^{-2} \text{ K}^{-1}$ also suggest a cloud feedback strength that is around the average of most climate models. The notable exception to the tendency of GSRMs to produce lower or average cloud feedbacks is that of Tsushima et al. (2014) who find that a large increase in high-clouds with warming results in a much higher cloud feedback than conventional GCMs. The results from SCREAM 3 km differ from most of these previous results by suggesting a more positive total feedback due to a strongly positive cloud feedback, mainly showing up in the shortwave. As the zonal-mean shortwave cloud feedback in Figure 6d shows, the shortwave cloud feedback is positive across most latitudes, barring some negative cloud feedback at high latitudes, which is still on the more positive end of what most CMIP models predict. SCREAM 3 km also does not show a relatively strong longwave cloud feedback that is found in NICAM experiments (Satoh et al., 2018).

The implications of the agreement and disagreement of the radiative feedbacks in SCREAM 3 km compared to CMIP-class models and existing GSRM studies are two-fold. First, the general agreement in the lapse-rate and relative humidity feedbacks in SCREAM 3 km with CMIP models suggests that predictions of these feedbacks are not largely skewed in climate models by the use of deep convective schemes. If they do disagree, their impact on the total feedback is small. In other words and perhaps not surprisingly, our results suggest that predicting the strength of the non-cloud, atmospheric feedbacks do not require expensive km-scale model simulations. Second, the strongly positive cloud feedback and the higher total feedback strength in SCREAM 3 km and the disagreement in the cloud feedbacks with existing GSRM studies shows that despite hopes that the removal of the deep convective scheme would reduce uncertainties in cloud feedbacks (Satoh et al., 2018), the spread in cloud feedbacks remains in GSRMs. Indeed, as the short survey of previous studies examining cloud feedbacks in the previous paragraph already shows, despite the suggestion of more negative cloud feedbacks in kilometer-scale models compared to GCMs, there is disagreement on the reasons for the negative cloud feedbacks between previous studies with NICAM, SAM, and SPCAM, indicating the diversity in cloud responses even among kilometer-scale models. Further supporting this, Webb et al. (2015) showed that the spread in feedback estimates are large even when the convection scheme is turned off in climate models. Deeper analysis into the type of cloud feedbacks that contribute to the total cloud feedback, such as those provided by Zelinka et al. (2022) and Chao et al. (2024), and further understanding of the processes that contribute to the positive cloud feedback in SCREAM 3 km will be necessary to assess the confidence we have in the cloud response.

Because these are computationally expensive models that very few modeling centers can perform, we have made an effort to save a wide range of outputs (as documented in Tables S2–S4 in Supporting Information S1) from the simulations to make available publicly at NERSC (Terai, 2024b). While questions related to variability or statistics of extreme events might not be answered in a pair of 1-year simulations, results from the CMIP model output and the five-member ensemble of SCREAM 12 km indicate that uncertainty in the feedbacks from analyzing just a pair of 1-year simulations are likely small compared to the intermodel spread and much can be learned from a 1-year simulation. When assessed with previous studies, the results from SCREAM 3 km show that there is a diversity of feedback responses even across GSRMs. Better understanding the root causes of that model diversity and their potential impacts on other predictions will become possible as the number of GSRMs performing such experiments grow. More than 35 years later, the experimental design proposed by Cess and Potter (1988) continues to help us better understand the climate response in today's state-of-the-science models.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Acknowledgments

This research was supported as part of the Energy Exascale Earth System Model (E3SM) project (<https://e3sm.org/>), funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research. The work of MDZ, LWC, and HM was supported by the U.S. Department of Energy (DOE) Regional and Global Model Analysis program area. HM and YZ are supported by the U.S. DOE Atmospheric System Research Program's Tying in High Resolution E3SM with ARM Data (THREAD) project. This paper was prepared under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. IM Release number LLNL-JRNL-2001459. Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC (NTESS), a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration (DOE/NNSA) under contract DE-NA0003525. This written work is authored by an employee of NTESS. The employee owns the right, title and interest in and to the written work and is responsible for its contents. Any subjective views or opinions that might be expressed in the written work do not necessarily represent the views of the U.S. Government. The publisher acknowledges that the U.S. Government retains a non-exclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this written work or allow others to do so, for U.S. Government purposes. The DOE will provide public access to results of federally sponsored research in accordance with the DOE Public Access Plan. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government. This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. This manuscript has been co-authored by Oak Ridge National Laboratory, operated by UT-Battelle, LLC under Contract No. DEAC05-00OR22725 with the U.S. Department of Energy. An award of computer time was also provided by the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program, performed using resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

Data Availability Statement

Instructions for accessing simulation data from the SCREAM simulations can be found at <https://zenodo.org/records/14579433> (Terai, 2024b), and the version of the code that was used to perform the simulations is archived at <https://zenodo.org/records/14578966> (Terai, 2024a).

CERES_EBAF_Ed4.2 data were obtained from the NASA Langley Research Center CERES ordering tool at <https://ceres.larc.nasa.gov/data/>. We also used the ERA5 monthly- (Hersbach et al., 2019a, 2019b), which can be downloaded from the Copernicus Climate Change Service (2023). Monthly averaged GPM-IMERG data (GPM_3IMERGM) were obtained from https://disc.gsfc.nasa.gov/datasets/GPM_3IMERGM_07/summary (Precipitation Processing System (PPS) At NASA GSFC, 2023). Monthly averaged GPCP v3.2 data can be obtained from https://disc.gsfc.nasa.gov/datasets/GPCPMON_3.2/summary (Huffman, 2021). The PyFLEXTRKR analysis packaged is available at <https://github.com/FlexTRKR/PyFLEXTRKR>.

References

- Bertagna, L., Guba, O., Taylor, M. A., Foucar, J. G., Larkin, J., Bradley, A. M., et al. (2020). A performance-portable nonhydrostatic atmospheric dycore for the energy exascale Earth system model running at cloud-resolving resolutions. In *Sc20: International conference for high performance computing, networking, storage and analysis* (pp. 1–14). <https://doi.org/10.1109/SC41405.2020.00096>
- Blossey, P. N., Bretherton, C. S., & Wyant, M. C. (2009). Subtropical low cloud response to a warmer climate in a superparameterized climate model. Part II: Column modeling with a cloud resolving model. *Journal of Advances in Modeling Earth Systems*, 1(3), 8. <https://doi.org/10.3894/JAMES.2009.1.8>
- Bogenschütz, P. A., Eldred, C., & Caldwell, P. M. (2023). Horizontal resolution sensitivity of the simple convection-permitting E3SM atmosphere model in a doubly-periodic configuration. *Journal of Advances in Modeling Earth Systems*, 15(7), e2022MS003466. <https://doi.org/10.1029/2022MS003466>
- Bogenschütz, P. A., & Krueger, S. K. (2013). A simplified PDF parameterization of subgrid-scale clouds and turbulence for cloud-resolving models. *Journal of Advances in Modeling Earth Systems*, 5(2), 195–211. <https://doi.org/10.1002/jame.20018>
- Bogenschütz, P. A., Zhang, Y., Zheng, X., Tian, Y., Zhang, M., Lin, L., et al. (2025). Exposing process-level biases in a global cloud permitting model with arm observations. *Journal of Geophysical Research: Atmospheres*, 130, e2024JD043059. <https://doi.org/10.1029/2024JD043059>
- Bretherton, C. S., & Khairoutdinov, M. F. (2015). Convective self-aggregation feedbacks in near-global cloud-resolving simulations of an aquaplanet. *Journal of Advances in Modeling Earth Systems*, 7(4), 1765–1787. <https://doi.org/10.1002/2015MS000499>
- Brient, F., & Bony, S. (2013). Interpretation of the positive low-cloud feedback predicted by a climate model under global warming. *Climate Dynamics*, 40(9), 2415–2431. <https://doi.org/10.1007/s00382-011-1279-7>
- Caldwell, P. M., Terai, C. R., Hillman, B., Keen, N. D., Bogenschütz, P., Lin, W., et al. (2021). Convection-permitting simulations with the E3SM global atmosphere model. *Journal of Advances in Modeling Earth Systems*, 13(11), e2021MS002544. <https://doi.org/10.1029/2021ms002544>
- Cess, R. D., & Potter, G. L. (1988). A methodology for understanding and intercomparing atmospheric climate feedback processes in general circulation models. *Journal of Geophysical Research*, 93(D7), 8305–8314. <https://doi.org/10.1029/JD093iD07p08305>
- Cess, R. D., Potter, G. L., Blanchet, J. P., Boer, G. J., Ghan, S. J., Kiehl, J. T., et al. (1989). Interpretation of cloud-climate feedback as produced by 14 atmospheric general circulation models. *Science*, 245(4917), 513–516. <https://doi.org/10.1126/science.245.4917.513>
- Chao, L., Zelinka, M. D., & Dessler, A. E. (2024). Evaluating Cloud feedback components in observations and their representation in climate models. *Journal of Geophysical Research: Atmospheres*, 129(2), e2023JD039427. <https://doi.org/10.1029/2023JD039427>
- Chinita, M. J., Witte, M., Kurowski, M. J., Teixeira, J., Suselj, K., Matheou, G., & Bogenschütz, P. (2023). Improving the representation of shallow cumulus convection with the simplified-higher-order-closure-mass-flux (SHOC+MF v1.0) approach. *Geoscientific Model Development*, 16(7), 1909–1924. <https://doi.org/10.5194/gmd-16-1909-2023>
- Christensen, M. W., Carrió, G. G., Stephens, G. L., & Cotton, W. R. (2013). Radiative impacts of free-tropospheric clouds on the properties of marine stratocumulus. *Journal of the Atmospheric Sciences*, 70(10), 3102–3118. <https://doi.org/10.1175/JAS-D-12-0287.1>
- Copernicus Climate Change Service. (2023). ERA5 monthly averaged data on pressure levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). <https://doi.org/10.24381/cds.6860a573>
- DeAngelis, A. M., Qu, X., Zelinka, M. D., & Hall, A. (2015). An observational radiative constraint on hydrologic cycle intensification. *Nature*, 528(7581), 249–253. <https://doi.org/10.1038/nature15770>
- Dennis, J., Edwards, J., Evans, K. J., Guba, O., Lauritzen, P., Mirin, A. A., et al. (2012). CAM-SE: A scalable spectral element dynamical core for the Community Atmosphere Model. *International Journal of High Performance Computing Applications*, 26(1), 74–89. <https://doi.org/10.1177/1094342011428142>
- Dennis, J., Fournier, A., Spitz, W. F., St-Cyr, A., Taylor, M. A., Thomas, S. J., & Tufo, H. (2005). High-resolution mesh convergence properties and parallel efficiency of a spectral element atmospheric dynamical core. *International Journal of High Performance Computing Applications*, 19(3), 225–235. <https://doi.org/10.1177/1094342005056108>
- Donahue, A. S., Caldwell, P. M., Bertagna, L., Beydoun, H., Bogenschütz, P. A., Bradley, A. M., et al. (2024). To exascale and beyond—The Simple Cloud-Resolving E3SM Atmosphere Model (SCREAM), a performance portable global atmosphere model for cloud-resolving scales. *Journal of Advances in Modeling Earth Systems*, 16(7), e2024MS004314. <https://doi.org/10.1029/2024MS004314>
- Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E., & Wimmer, W. (2012). The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system. *Remote Sensing of Environment*, 116, 140–158. <https://doi.org/10.1016/j.rse.2010.10.017>
- Feng, Z., Hardin, J., Barnes, H. C., Li, J., Leung, L. R., Varble, A., & Zhang, Z. (2023). PyFLEXTRKR: A flexible feature tracking Python software for convective cloud analysis. *Geoscientific Model Development*, 16(10), 2753–2776. <https://doi.org/10.5194/gmd-16-2753-2023>
- Garcia, R. R., Smith, A. K., Kinnison, D. E., Cámara, d. I., & Murphy, D. J. (2017). Modification of the gravity wave parameterization in the whole atmosphere community climate model: Motivation and results. *Journal of the Atmospheric Sciences*, 74(1), 275–291. <https://doi.org/10.1175/JAS-D-16-0104.1>
- Gregory, J. M., & Andrews, T. (2016). Variation in climate sensitivity and feedback parameters during the historical period. *Geophysical Research Letters*, 43(8), 3911–3920. <https://doi.org/10.1002/2016GL068406>

Pacific Northwest National Laboratory is operated by Battelle for the U.S. Department of Energy under Contract DE-AC05-76RLO1830. The work at BNL was supported by the Energy Exascale Earth System Model (E3SM) project funded by Department of Energy under contract DE-SC0012704. Argonne National Laboratory's work was supported by the U.S. Department of Energy, Assistant Secretary for Environmental Management, Office of Science, under contract DE-AC02-06CH11357. We would like to thank Steve Klein for insightful discussions on early results and thank three anonymous reviewers for comments which have substantially improved the quality of the manuscript. We also acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP5 and CMIP6. We thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP and ESGF.

- Gregory, J. M., Ingram, W. J., Palmer, M. A., Jones, G. S., Stott, P. A., Thorpe, R. B., et al. (2004). A new method for diagnosing radiative forcing and climate sensitivity. *Geophysical Research Letters*, 31(3), L03205. <https://doi.org/10.1029/2003GL018747>
- Guendelman, I., Merlis, T. M., Cheng, K., Harris, L. M., Bretherton, C. S., Bolot, M., et al. (2024). The precipitation response to warming and CO₂ increase: A comparison of a global storm resolving model and CMIP6 models. *Geophysical Research Letters*, 51(7), e2023GL107008. <https://doi.org/10.1029/2023GL107008>
- Heim, C., Hentgen, L., Ban, N., & Schär, C. (2021). Inter-model variability in convection-resolving simulations of subtropical marine low clouds. *Journal of the Meteorological Society of Japan. Ser. II*, 99(5), 1271–1295. <https://doi.org/10.2151/jmsj.2021-062>
- Held, I. M., & Shell, K. M. (2012). Using relative humidity as a state variable in climate feedback analysis. *Journal of Climate*, 25(8), 2578–2582. <https://doi.org/10.1175/JCLI-D-11-00721.1>
- Held, I. M., & Soden, B. J. (2006). Robust responses of the hydrological cycle to global warming. *Journal of Climate*, 19(21), 5686–5699. <https://doi.org/10.1175/JCLI3990.1>
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. <https://doi.org/10.1002/qj.3803>
- Hersbach, H., Bell, B., Berrisford, P., Horányi, A., Muñoz-Sabater, J., Nicolas, J., et al. (2019a). ERA5 monthly averaged data on pressure levels from 1940 to present. *Copernicus Climate Change Service*. <https://doi.org/10.24381/CDS.6860A573>
- Hersbach, H., Bell, B., Berrisford, P., Horányi, A., Muñoz-Sabater, J., Nicolas, J., et al. (2019b). ERA5 monthly averaged data on single levels from 1940 to present. *Copernicus Climate Change Service*. <https://doi.org/10.24381/CDS.F17050D7>
- Hohenegger, C., Kornbluth, L., Klocke, D., Becker, T., Cioni, G., Engels, J. F., et al. (2020). Climate statistics in global simulations of the atmosphere, from 80 to 2.5 km Grid Spacing. *Journal of the Meteorological Society of Japan. Ser. II*, 98(1), 73–91. <https://doi.org/10.2151/jmsj.2020-005>
- Huang, Y., Xia, Y., & Tan, X. (2017). On the pattern of CO₂ radiative forcing and poleward energy transport. *Journal of Geophysical Research: Atmospheres*, 122(20), 10578–10593. <https://doi.org/10.1002/2017JD027221>
- Huffman, G. J. (2021). *GPCP precipitation level 3 monthly 0.5-degree V3.2*. NASA Goddard Earth Sciences Data and Information Services Center. <https://doi.org/10.5067/MEASURES/GPCP/DATA304>
- Huffman, G. J., Adler, R. F., Behrangi, A., Bolvin, D. T., Nelkin, E., & Song, Y. (2020). Algorithm Theoretical Basis Document (ATBD) for global precipitation climatology project version 3.1 precipitation data. Retrieved from https://docsserver.gesdisc.eosdis.nasa.gov/public/project/MEASURES/GPCP/GPCP_ATBD_V3.1.pdf
- Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K.-L., Joyce, R. J., Kidd, C., et al. (2020). Integrated Multi-satellite Retrievals for the Global Precipitation Measurement (GPM) Mission (IMERG). In V. Levizzani, C. Kidd, D. B. Kirschbaum, C. D. Kummerow, K. Nakamura, & F. J. Turk (Eds.), *Satellite Precipitation Measurement* (Vol. 1, pp. 343–353). Springer International Publishing. https://doi.org/10.1007/978-3-030-24568-9_19
- Kang, S. M., & Polvani, L. M. (2011). The interannual relationship between the latitude of the eddy-driven jet and the edge of the Hadley cell. *Journal of Climate*, 24(2), 563–568. <https://doi.org/10.1175/2010JCLI4077.1>
- Kato, S., Rose, F. G., Rutan, D. A., Thorsen, T. J., Loeb, N. G., Doelling, D. R., et al. (2018). Surface Irradiances of Edition 4.0 Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) data product. *Journal of Climate*, 31(11), 4501–4527. <https://doi.org/10.1175/JCLI-D-17-0523.1>
- Kiehl, J. T., Shields, C. A., Hack, J. J., & Collins, W. D. (2006). The climate sensitivity of the community climate system model version 3 (CCSM3). *Journal of Climate*, 19(11), 2584–2596. <https://doi.org/10.1175/JCLI3747.1>
- Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J. G., et al. (2018). 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(2), 895–918. <https://doi.org/10.1175/JCLI-D-17-0208.1>
- Lucas, C., Timbal, B., & Nguyen, H. (2014). The expanding tropics: A critical assessment of the observational and modeling studies. *WIREs Climate Change*, 5(1), 89–112. <https://doi.org/10.1002/wcc.251>
- Mahfouz, N., Beydoun, H., Mülmenstädt, J., Keen, N., Varble, A. C., Bertagna, L., et al. (2025). Prescribing the aerosol effective radiative forcing in the simple cloud-resolving E3SM atmosphere model v1. *EGU sphere*, 2025, 1–25. <https://doi.org/10.5194/egusphere-2025-1868>
- Manabe, S., & Stouffer, R. J. (1980). Sensitivity of a global climate model to an increase of CO₂ concentration in the atmosphere. *Journal of Geophysical Research*, 85(C10), 5529–5554. <https://doi.org/10.1029/JC085iC10p05529>
- Medeiros, B., & Stevens, B. (2011). Revealing differences in GCM representations of low clouds. *Climate Dynamics*, 36(1–2), 385–399. <https://doi.org/10.1007/s00382-009-0694-5>
- Merlis, T. M., Cheng, K.-Y., Guendelman, I., Harris, L., Bretherton, C. S., Bolot, M., et al. (2024). Climate sensitivity and relative humidity changes in global storm-resolving model simulations of climate change. *Science Advances*, 10(26), eadn5217. <https://doi.org/10.1126/sciadv.adn5217>
- Miura, H., Tomita, H., Nasuno, T., Iga, S., Satoh, M., & Matsuno, T. (2005). A climate sensitivity test using a global cloud resolving model under an aqua planet condition. *Geophysical Research Letters*, 32(19), 2005GL023672. <https://doi.org/10.1029/2005GL023672>
- Morrison, H., & Milbrandt, J. A. (2015). Parameterization of cloud microphysics based on the prediction of bulk ice particle properties. Part I: Scheme description and idealized tests. *Journal of the Atmospheric Sciences*, 72(1), 287–311. <https://doi.org/10.1175/JAS-D-14-0065.1>
- Narenpitak, P., Bretherton, C. S., & Khairoutdinov, M. F. (2017). Cloud and circulation feedbacks in a near-global aquaplanet cloud-resolving model. *Journal of Advances in Modeling Earth Systems*, 9(2), 1069–1090. <https://doi.org/10.1002/2016MS000872>
- Nowak, J. L., Dragaud, I. C. D. V., Lee, J., Dziekan, P., Mellado, J. P., & Stevens, B. (2025). A first look at the global climatology of low-level clouds in storm resolving models. *Journal of Advances in Modeling Earth Systems*, 17(3), e2024MS004340. <https://doi.org/10.1029/2024MS004340>
- Pincus, R., Mlawer, E. J., & Delamere, J. S. (2019). Balancing accuracy, efficiency, and flexibility in radiation calculations for dynamical models. *Journal of Advances in Modeling Earth Systems*, 11(10), 3074–3089. <https://doi.org/10.1029/2019MS001621>
- Precipitation Processing System (PPS) At NASA GSFC. (2023). *GPM IMERG final precipitation L3 1 month 0.1 degree x 0.1 degree V07*. NASA Goddard Earth Sciences Data and Information Services Center. <https://doi.org/10.5067/GPM/IMERG/3B-MONTH/07>
- Prein, A. F., Liu, C., Ikeda, K., Trier, S. B., Rasmussen, R. M., Holland, G. J., & Clark, M. P. (2017). Increased rainfall volume from future convective storms in the US. *Nature Climate Change*, 7(12), 880–884. <https://doi.org/10.1038/s41558-017-0007-7>
- Prein, A. F., Rasmussen, R. M., Wang, D., & Giangrande, S. E. (2021). Sensitivity of organized convective storms to model grid spacing in current and future climates. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2195), 20190546. <https://doi.org/10.1098/rsta.2019.0546>
- Qin, Y., Zelinka, M. D., & Klein, S. A. (2022). On the correspondence between atmosphere-only and coupled simulations for radiative feedbacks and forcing from CO₂. *Journal of Geophysical Research: Atmospheres*, 127(3), e2021JD035460. <https://doi.org/10.1029/2021JD035460>

- Qin, Y., Zheng, X., Klein, S. A., Zelinka, M. D., Ma, P., Golaz, J., & Xie, S. (2024). Causes of reduced climate sensitivity in E3SM from version 1 to version 2. *Journal of Advances in Modeling Earth Systems*, 16(1), e2023MS003875. <https://doi.org/10.1029/2023MS003875>
- Qu, X., Hall, A., Klein, S. A., & Caldwell, P. M. (2014). On the spread of changes in marine low cloud cover in climate model simulations of the 21st century. *Climate Dynamics*, 42(9–10), 2603–2626. <https://doi.org/10.1007/s00382-013-1945-z>
- Reeves Eyre, J. E. J., Zeng, X., & Zhang, K. (2021). Ocean surface flux algorithm effects on earth system model energy and water cycles. *Frontiers in Marine Science*, 8, 642804. <https://doi.org/10.3389/fmars.2021.642804>
- Ringer, M. A., Andrews, T., & Webb, M. J. (2014). Global-mean radiative feedbacks and forcing in atmosphere-only and coupled atmosphere-ocean climate change experiments. *Geophysical Research Letters*, 41(11), 4035–4042. <https://doi.org/10.1002/2014GL060347>
- Satoh, M., Noda, A. T., Seiki, T., Chen, Y.-W., Kodama, C., Yamada, Y., et al. (2018). Toward reduction of the uncertainties in climate sensitivity due to cloud processes using a global non-hydrostatic atmospheric model. *Progress in Earth and Planetary Science*, 5(1), 67. <https://doi.org/10.1186/s40645-018-0226-1>
- Schiro, K. A., Su, H., Ahmed, F., Dai, N., Singer, C. E., Gentine, P., et al. (2022). Model spread in tropical low cloud feedback tied to overturning circulation response to warming. *Nature Communications*, 13(1), 7119. <https://doi.org/10.1038/s41467-022-34787-4>
- Segura, H., Pedruzo-Bagazgoitia, X., Weiss, P., Müller, S. K., Rackow, T., Lee, J., et al. (2025). nextGEMS: Entering the era of kilometer-scale Earth system modeling. *EGUsphere*, 1–39. <https://doi.org/10.5194/egusphere-2025-509>
- Shell, K. M., Kiehl, J. T., & Shields, C. A. (2008). Using the radiative kernel technique to calculate climate feedbacks in NCAR's community atmospheric model. *Journal of Climate*, 21(10), 2269–2282. <https://doi.org/10.1175/2007JCLI2044.1>
- Sherwood, S. C., Bony, S., & Dufresne, J.-L. (2014). Spread in model climate sensitivity traced to atmospheric convective mixing. *Nature*, 505(7481), 37–42. <https://doi.org/10.1038/nature12829>
- Sherwood, S. C., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., et al. (2020). An assessment of Earth's climate sensitivity using multiple lines of evidence. *Reviews of Geophysics*, 58(4), e2019RG000678. <https://doi.org/10.1029/2019RG000678>
- Soden, B. J., Held, I. M., Colman, R., Shell, K. M., Kiehl, J. T., & Shields, C. A. (2008). Quantifying climate feedbacks using radiative kernels. *Journal of Climate*, 21(14), 3504–3520. <https://doi.org/10.1175/2007JCLI2110.1>
- Stevens, B., Satoh, M., Auger, L., Biercamp, J., Bretherton, C. S., Chen, X., et al. (2019). DYAMOND: The DYNAMics of the atmospheric general circulation modeled on non-hydrostatic domains. *Progress in Earth and Planetary Science*, 6(1), 61. <https://doi.org/10.1186/s40645-019-0304-z>
- Takasuka, D., Kodama, C., Suematsu, T., Ohno, T., Yamada, Y., Seiki, T., et al. (2024). How can we improve the seamless representation of climatological statistics and weather toward reliable global K-scale climate simulations? *Journal of Advances in Modeling Earth Systems*, 16(2), e2023MS003701. <https://doi.org/10.1029/2023MS003701>
- Takasuka, D., Satoh, M., Miyakawa, T., Kodama, C., Klocke, D., Stevens, B., et al. (2024). A protocol and analysis of year-long simulations of global storm-resolving models and beyond. <https://doi.org/10.21203/rs.3.rs-4458164/v1>
- Taylor, M. A., Guba, O., Steyer, A., Ullrich, P. A., Hall, D. M., & Eldrid, C. (2020). An energy consistent discretization of the nonhydrostatic equations in primitive variables. *Journal of Advances in Modeling Earth Systems*, 12(1), e2019MS001783. <https://doi.org/10.1029/2019ms001783>
- Terai, C. (2024a). Model code used for Cess-Potter experiments with SCREAMv1. *Zenodo*. <https://doi.org/10.5281/zenodo.14578966>
- Terai, C. (2024b). Simulation output and input data from Cess-Potter experiments with SCREAMv1. *Zenodo*. <https://doi.org/10.5281/zenodo.14579433>
- Trott, C. R., Lebrun-Grandié, D., Arndt, D., Ciesko, J., Dang, V., Ellingwood, N., et al. (2022). Kokkos 3: Programming model extensions for the exascale era. *IEEE Transactions on Parallel and Distributed Systems*, 33(4), 805–817. <https://doi.org/10.1109/TPDS.2021.3097283>
- Tsushima, Y., Iga, S.-i., Tomita, H., Satoh, M., Noda, A. T., & Webb, M. J. (2014). High cloud increase in a perturbed SST experiment with a global nonhydrostatic model including explicit convective processes. *Journal of Advances in Modeling Earth Systems*, 6(3), 571–585. <https://doi.org/10.1002/2013MS000301>
- Vecchi, G. A., & Soden, B. J. (2007). Global warming and the weakening of the tropical circulation. *Journal of Climate*, 20(17), 4316–4340. <https://doi.org/10.1175/JCLI4258.1>
- Viovy, N. (2018). Cruncep version 7—Atmospheric forcing data for the community land model. *Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory*. <https://doi.org/10.5065/PZ8F-F017>
- Vogel, R., Albright, A. L., Vial, J., George, G., Stevens, B., & Bony, S. (2022). Strong cloud–circulation coupling explains weak trade cumulus feedback. *Nature*, 612(7941), 696–700. <https://doi.org/10.1038/s41586-022-05364-y>
- Webb, M. J., Andrews, T., Bodas-Salcedo, A., Bony, S., Bretherton, C. S., Chadwick, R., et al. (2017). The Cloud Feedback Model Intercomparison Project (CFMIP) contribution to CMIP6. *Geoscientific Model Development*, 10(1), 359–384. <https://doi.org/10.5194/gmd-10-359-2017>
- Webb, M. J., Lock, A. P., Bretherton, C. S., Bony, S., Cole, J. N. S., Idelkadi, A., et al. (2015). The impact of parametrized convection on cloud feedback. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 373(2054), 20140414. <https://doi.org/10.1098/rsta.2014.0414>
- Webb, M. J., Lock, A. P., & Ogura, T. (2024). What are the main causes of positive subtropical low cloud feedbacks in climate models? *Journal of Advances in Modeling Earth Systems*, 16(1), e2023MS003716. <https://doi.org/10.1029/2023MS003716>
- Wu, W., Ma, H., Lafferty, D. C., Feng, Z., Ullrich, P., Tang, Q., et al. (2024). Assessment of storm-associated precipitation and its extremes using observational data sets and climate model short-range hindcasts. *Journal of Geophysical Research: Atmospheres*, 129(9), e2023JD039697. <https://doi.org/10.1029/2023JD039697>
- Wyant, M. C., Bretherton, C. S., & Blossey, P. N. (2009). Subtropical low cloud response to a warmer climate in a superparameterized climate model. Part I: Regime sorting and physical mechanisms. *Journal of Advances in Modeling Earth Systems*, 1(3), 7. <https://doi.org/10.3894/JAMES.2009.1.7>
- Wyant, M. C., Khairoutdinov, M., & Bretherton, C. S. (2006). Climate sensitivity and cloud response of a GCM with a superparameterization. *Geophysical Research Letters*, 33(6), 2005GL025464. <https://doi.org/10.1029/2005GL025464>
- Zelinka, M. D., Klein, S. A., Qin, Y., & Myers, T. A. (2022). Evaluating climate models' cloud feedbacks against expert judgment. *Journal of Geophysical Research: Atmospheres*, 127(2), e2021JD035198. <https://doi.org/10.1029/2021JD035198>
- Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., et al. (2020). Causes of higher climate sensitivity in CMIP6 models. *Geophysical Research Letters*, 47(1), e2019GL085782. <https://doi.org/10.1029/2019GL085782>
- Zhang, M., & Bretherton, C. (2008). Mechanisms of low cloud–climate feedback in idealized single-column simulations with the community atmospheric model, version 3 (CAM3). *Journal of Climate*, 21(18), 4859–4878. <https://doi.org/10.1175/2008JCLI2237.1>

- Zhao, M. (2022). A study of AR-TS-and MCS-associated precipitation and extreme precipitation in present and warmer climates. *Journal of Climate*, 35(2), 479–497. <https://doi.org/10.1175/JCLI-D-21-0145.1>
- Zhou, C., Zelinka, M. D., & Klein, S. A. (2017). Analyzing the dependence of global cloud feedback on the spatial pattern of sea surface temperature change with a Green's function approach. *Journal of Advances in Modeling Earth Systems*, 9(5), 2174–2189. <https://doi.org/10.1002/2017MS001096>