Two-moment bulk cloud microphysics with prognostic precipitation in the GFDL CM4.0 model: Performance and simulation characteristics

Huan Guo[1], Yi Ming[2], Song-Miao Fan[3], Andrew T. Wittenberg[4], Rong Zhang[5], Ming Zhao[5], and Linjiong Zhou[2]

[1]NOAA/GFDL
[2]Geophysical Fluid Dynamics Laboratory
[3]National Oceanic and Atmospheric Administration (NOAA)
[4]NOAA Geophysical Fluid Dynamics Laboratory
[5]GFDL/NOAA

November 26, 2022

Abstract

We describe the model performance and simulation characteristics of a new global coupled climate model configuration, CM4-MG2. Beginning with the Geophysical Fluid Dynamics Laboratory's fourth-generation physical climate model (CM4.0), we incorporate a two-moment Morrison-Gettelman bulk cloud microphysics scheme with prognostic precipitation (MG2), and a mineral dust and temperature-dependent cloud ice nucleation scheme. We then conduct and analyze a set of fully coupled atmosphere-ocean-land-sea ice simulations, following Coupled Model Intercomparison Project Phase 6 (CMIP6) protocols. CM4-MG2 generally captures CM4.0's baseline simulation characteristics, but with several improvements, including better marine stratocumulus clouds off the west coasts of Africa and North and South America, a reduced bias toward "double" Intertropical Convergence Zones south of the equator, and a stronger Madden-Julian Oscillation (MJO). Some degraded features are also identified, including excessive Arctic sea ice extent and a stronger-than-observed El Nino-Southern Oscillation (ENSO). Compared to CM4.0, the climate sensitivity is reduced by about 10\% in CM4-MG2.

Two-moment bulk cloud microphysics with prognostic precipitation in the GFDL CM4.0 model: Performance and simulation characteristics

Huan Guo[1], Yi Ming[1], Songmiao Fan[1], Andrew T. Wittenberg[1], Rong Zhang[1], Ming Zhao[1], Linjiong Zhou[2]

.

[1]NOAA/OAR/Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA
[2]Cooperative Institute for Modeling the Earth System, Program in Oceanic and Atmospheric Sciences, Princeton University, Princeton, NJ, USA

Key Points:

- Two-moment cloud microphysics with prognostic precipitation, and an ice nucleation scheme, have been incorporated in CM4.0, named CM4-MG2.
- The overall performance of CM4-MG2 is comparable to or better than CM4.0.
- Notable improvements include enhanced coastal stratocumulus and a stronger MJO. CM4-MG2 also has a lower climate sensitivity than CM4.0.

Corresponding author: Huan Guo, `huan.guo@noaa.gov`

Abstract

We describe the model performance and simulation characteristics of a new global coupled climate model configuration, CM4-MG2. Beginning with the Geophysical Fluid Dynamics Laboratory's fourth-generation physical climate model (CM4.0), we incorporate a two-moment Morrison-Gettelman bulk cloud microphysics scheme with prognostic precipitation (MG2), and a mineral dust and temperature-dependent cloud ice nucleation scheme. We then conduct and analyze a set of fully coupled atmosphere-ocean-land-sea ice simulations, following Coupled Model Intercomparison Project Phase 6 (CMIP6) protocols. CM4-MG2 generally captures CM4.0's baseline simulation characteristics, but with several improvements, including better marine stratocumulus clouds off the west coasts of Africa and North and South America, a reduced bias toward "double" Intertropical Convergence Zones south of the equator, and a stronger Madden-Julian Oscillation (MJO). Some degraded features are also identified, including excessive Arctic sea ice extent and a stronger-than-observed El Niño-Southern Oscillation (ENSO). Compared to CM4.0, the climate sensitivity is reduced by about 10% in CM4-MG2.

Plain Language Summary

A sophisticated cloud microphysical scheme, along with a mineral dust and temperature-dependent ice nucleation scheme, have been implemented in a new configuration of the Geophysical Fluid Dynamics Laboratory's most recent climate model (CM4.0). This microphysical scheme predicts both mass and number concentrations of cloud drops, ice crystals, rain, and snow, and treats aerosol-cloud interactions more consistently. The ice nucleation on mineral dust aerosol in large-scale clouds is represented more realistically. Centennial-scale global coupled atmosphere-ocean-land-sea ice simulations from this new configuration compare favorably with observations — with improved subtropical stratocumulus clouds and better tropical intraseasonal variability (i.e. the 30- to 90-day Madden-Julian Oscillation). The new configuration also reduces the magnitude of future global warming in response to anthropogenic emissions.

1 Introduction

CM4.0 is the Geophysical Fluid Dynamics Laboratory (GFDL)'s fourth-generation physical general circulation model (GCM), participating in the Coupled Model Intercomparison Project Phase 6 (CMIP6) (Eyring et al., 2016). With successive upgrades in its model components (M. Zhao et al., 2018a, 2018b; Adcroft et al., 2019; Milly et al., 2014; Shevliakova et al., 2009), CM4.0 presents high-fidelity simulations of top-of-atmosphere (TOA) radiative fluxes, mean atmospheric state, Intertropical Convergence Zone (ITCZ), El Niño-Southern Oscillation (ENSO), ocean boundary currents, among others (Held et al., 2019).

However, there have been limited upgrades in the cloud microphysics of the atmospheric component since GFDL's second-generation model (CM2) (Delworth et al., 2006). The cloud microphysics parameterization in CM4.0 is the Rotstayn-Klein (RK) scheme, which is a one-moment^{+} bulk scheme with diagnostic precipitation (L. D. Rotstayn, 1997; L. Rotstayn et al., 2000; Jakob & Klein, 2000; GFDL Global Atmosphere Model Development Team, 2004; Donner et al., 2011; Golaz et al., 2011). The RK scheme prognoses the mass mixing ratios of cloud water and ice as well as cloud droplet number concentration (i.e., one-moment^{+} or partially two-moment), while the mass mixing ratios of rain and snow are diagnosed. The diagnostic precipitation treatment is efficient computationally, but there are a few issues. First, it distorts the relative importance of autoconversion and accretion for rain formation. Rain is diagnosed and removed in a single model time step, artificially suppressing accretion that depends on existing rain water and shifting the rain formation towards autoconversion. This does not conform to the observational constraint on the process level (Gettelman et al., 2013, 2015b). Second, the bias towards autoconversion likely amplifies aerosol indirect effects, because the autoconversion strongly depends on droplet size distribution and/or number concentration. The overestimate of autoconversion is one reason why the response of liquid water path (LWP) to aerosols is too strong (positive) in many GCMs (Quaas et al., 2009; M. Wang et al., 2012). Recent satellite observations and global cloud-resolving model simulations have also suggested that aerosol indirect effects might have been overestimated because the response of LWP to aerosols could be either positive or negative or neutral (Sato et al., 2018; Toll et al., 2019). Third, the neglect of precipitation advection is problematic in high-resolution atmospheric models. For example, given a 10 m s^{-1} horizontal wind speed and a 1 m s^{-1} fall velocity, falling 2 km means that precipitation has been advected to another grid box for horizontal grid spacing finer than 20 km. Hence the advection of precipitation is important as model resolution becomes more and more refined.

Furthermore, the RK scheme does not treat ice crystal number concentration (N_i) explicitly. Instead, it approximates N_i based on Meyers et al. (1992) in parameterizing Wegener–Bergeron–Findeisen (WBF) process. The concerns about the Meyers scheme are mainly two-fold in GFDL's Atmosphere Model version 4.0 (AM4.0). First, the Meyers ice nucleation scheme depends on temperature or ice supersaturation, not on aerosols. Hence the aerosol effects on ice clouds are missing. Second, the annual mean N_i estimated with the Meyers scheme is likely biased high, leading to a fast WBF conversion of supercooled liquid to ice. As a result, the supercooled liquid cloud fraction in the mixed-phase cloud regime is biased low when compared to satellite observations (Fan et al., 2019). As pinpointed by Tan et al. (2016), the supercooled fraction is closely linked to cloud-phase feedback via glaciation rate, and thus impacts the estimate of climate sensitivity and the fidelity of current and future climate simulations.

In order to address these issues and represent the aerosol indirect effects more realistically, Guo et al. (2021) implemented the two-moment Morrison-Gettelman cloud microphysics with prognostic precipitation (MG2 hereafter) and a temperature- and dust-dependent ice nucleation scheme into AM4.0 (Gettelman & Morrison, 2015a; Gettelman et al., 2015b; Fan et al., 2017, 2019). This configuration is termed AM4-MG2. MG2 is a bulk scheme by assuming that cloud particles follow a gamma distribution. It explicitly predicts the mass mixing ratios and number concentrations (two moments) of cloud water, ice, rain, and snow. Therefore, it is expected to treat the aerosol-cloud interactions more consistently. Moreover, the temperature- and dust-dependent ice nucleation parameterization is obtained by fitting air parcel model results, which agree well with laboratory experiments and in situ aircraft measurements. The air parcel model considers deposition nucleation, condensation nucleation, and immersion freezing on mineral dust particles. It turns out that AM4-MG2 simulations show weaker (less negative) aerosol radiative effects, more realistic supercooled liquid fraction, and improved stratocumulus clouds.

As a follow-up, we have applied the AM4-MG2 configuration under the coupled model framework of CM4.0, referred to as CM4-MG2. This paper aims to document the model performance and simulation characteristics of CM4-MG2. We give brief descriptions of the model components in Section 2. Section 3 discusses the CM4-MG2 fully coupled atmosphere-ocean-land-sea ice global simulation results, including pre-industrial control simulation, model mean climate of recent decades (1980-2014), climate variability, the twentieth century warming, and climate sensitivity and cloud feedback, as well as comparison to the base model CM4.0. Finally, a summary of results is given in Section 4.

2 Model description of CM4-MG2

2.1 Atmospheric component

The atmospheric component of CM4-MG2 is based on AM4.0 (M. Zhao et al., 2018b, 2018a). It uses the hydrostatic version of the GFDL Finite-Volume Cubed-Sphere (FV3) Dynamical Core (FV3) (Lin, 2004; Putman & Lin, 2007; L. Harris et al., 2020). The long-wave radiation code adopts the simplified exchange approximation (SEA) with updated spectral information and inclusion of CO_2 10 μm band (Fels & Schwarzkopf, 1975; Schwarzkopf & Fels, 1991). The shortwave code employs the 18-band formulation with updated H_2O, CO_2 and O_2 formulations and inclusion of the shortwave water vapor continuum and CH_4 and N_2O absorption (Freidenreich & Ramaswamy, 2005; Paynter & Ramaswamy, 2012, 2014). With these updates, the shortwave absorption error is reduced down to 1% within the line-by-line benchmark calculation (M. Zhao et al., 2018a). Both shallow convection and deep convection are uniformly treated by a "double-plume" scheme (M. Zhao et al., 2016, 2018b). Orographic gravity wave drag parameterization allows for arbitrary topography and considers nonlinear effects (Garner, 2005, 2018). Nonorographic gravity wave drag is parameterized following Alexander and Dunkerton (1999). The turbulent diffusivities in the planetary boundary layer (PBL) are parameterized following Lock et al. (2000). The large-scale cloud fraction (or cloud macrophysics) is prognosed according to Tiedtke (1993). The bulk aerosol scheme, including 17 transported aerosol tracers, is similar to that in GFDL's Atmosphere Model version 3 (AM3) (Donner et al., 2011), but with a "light" chemistry that turns off photochemistry and stratospheric chemistry (M. Zhao et al., 2018a; Salzmann et al., 2010). Aerosols are simulated from emissions using prescribed ozone and other oxidants (e.g., OH), and are linked to the cloud microphysics through the parameterization of droplet activation. The droplet activation depends on aerosol mass, chemical composition, and vertical velocity, following the parameterization detailed in Ming et al. (2006, 2007). Important changes in the atmospheric component from CM4.0 to CM4-MG2 include:

1. the replacement of the RK cloud microphysics with the MG2 microphysics.

2. the incorporation of the mineral dust and temperature-dependent ice nucleation parameterization.

3. the inclusion of rain and snow radiation effects. The shortwave radiative properties of rain are based on the Mie theory (Savijarvi, 1997), while the shortwave radiative properties of snow are parameterized following Fu et al. (1995). The longwave properties of rain and snow are derived assuming that rain and snow are spherical particles.

More details on the atmospheric component of CM4-MG2 are available in the AM4-MG2 documentation paper by Guo et al. (2021).

2.2 Land, ocean, and sea ice components

The remaining components in CM4-MG2 are identical to those in CM4.0 (Held et al., 2019). The land component is referred to as LM4.0.1, which is similar to LM4.0 as documented in M. Zhao et al. (2018a, 2018b) but with dynamic vegetation, enhanced snow-covered glacial albedo, and tiling structure interacting with atmosphere. The ocean component: OM4p25, is described in Adcroft et al. (2019). It uses a hybrid depth-isopycnal coordinate (Bleck, 2002; Adcroft & Hallberg, 2006), and about 25 km horizontal resolution without mesoscale eddy parameterization. The sea ice component adopts the Sea Ice Simulator version 2 (SIS2) (Adcroft et al., 2019), which is based upon the earlier sea ice model version employed in CM2 (Delworth et al., 2006). But the code was completely rewritten and contains many ice physics changes. SIS2 shares the same horizontal grid layouts (i.e., the Arakawa C-grid) as OM4p25, but with four sea ice layers and one snow layer vertically. There are 5 sea ice thickness categories bounded at 0.1, 0.3, 0.7, and 1.1 m. The thinnest category extends down to zero and the thickest is unbounded. These 5 categories are concentrated in the low sea ice thickness categories, because of the lack of a subgrid ice ridging scheme (Adcroft et al., 2019).

3 Model simulations and results

With CM4-MG2, we have conducted a suite of fully coupled atmosphere-ocean-land-sea ice CMIP6 Diagnosis, Evaluation, and Characterization of Klima (DECK) and historical simulations (Eyring et al., 2016), including 500-year pre-industrial control (piControl), 150-year CO_2 concentration increasing 1% per year (1pctCO2), 150-year abruptly quadrupled CO_2 (abrupt-4XCO2), and three historical ensemble (1850-2014) simulations (Table 1). The piControl experiment was initialized from the piControl spinup run at year 151, and was driven by the fixed forcing levels at 1850. The piControl spinup follows the same procedure as in CM4.0 where atmosphere and land states were from a 700-year piControl simulation with prototype configurations, and ocean and sea ice were based on the World Ocean Atlas January climatology (Antonov et al., 2006; Locarnini et al., 2006; Held et al., 2019). The 1pctCO2 and abrupt-4XCO2 experiments were branched off the piControl at year 101. The three historical ensemble simulations share the same ocean, sea ice, and land initial conditions spun off the piControl at year 101; but differ in the atmosphere initial condition which came from the piControl restart files at year 101, 140, and 182, respectively. Note that in order to have an apple-to-apple comparison, the CM4-MG2 fully coupled simulations were configured as close as possible to the CM4.0 simulations (Held et al., 2019).

All coupled simulations discussed in this study were run at nominal 1.0° horizontal resolution (or about 100 km) for atmosphere and land, and 0.25° horizontal resolution for ocean and sea ice. The atmospheric component uses 33 levels with a relatively "low top" of about 1 hPa in the vertical, while the ocean component has 75 vertical layers with about 2 m vertical spacing near the ocean surface and 250 m below 5000 m. The atmosphere physics time step is 30 min, and ocean baroclinic and barotropic time steps are 15 min and about 19 sec, respectively. The coupling frequency for all components is every 30 min.

3.1 Pre-industrial Control Experiment

In this section we will discuss the global-scale evolution of the CM4-MG2 piControl simulation. Figure 1(a) provides the net downward radiative flux at TOA. The TOA radiative flux generally fluctuates between -1.0 and +1.0 W m^{-2} with little model drift. Its 500-yr average is about 0.22 W m^{-2}. We also calculate the net heat flux out of the atmosphere at the surface, which is stable with an average of about 0.17 W m^{-2} over the 500-yr period. The non-zero difference between the TOA and surface fluxes suggests an artificial energy sink of 0.05 W m^{-2} in the CM4-MG2 model atmosphere (vs. 0.08 W m^{-2} in CM4.0). This sink stems from the inconsistent definitions of energy conservation between model dynamics and physics. For example, the atmospheric dynamic core considers the heat capacity for the total air (including condensed water) and the temperature dependence of latent heat, but atmospheric physics does not (Lin, 2004; Putman & Lin, 2007; Yano & Maarten, 2017; Zhou et al., 2019). An energy fix term for this inconsistency has been introduced in the AMIP (Atmospheric Model Intercomparison Project) mode, but gives rise to an energy sink (or imbalance) in the fully coupled mode (Held et al., 2019). However, the energy imbalance here is small relative to the radiative forcing caused by anthropogenic emissions, so we do not expect it would impose significant impacts on the model climate (Golaz et al., 2019).

Figures 1(c) and (d) present the time evolution of the global mean surface air temperature at 2 m (T_{air}) and sea surface temperature (SST) from the OM4p25 outputs. Both T_{air} and SST show slightly warming trends (+0.018°C/century for T_{air} and +0.015°C/century for SST). This is partly associated with the Southern Ocean that has not reached the equilibrium or steady state in the CM4-MG2 piControl, similar to what is reported in Held et al. (2019). Compared to the HadISST over 1880-1900 (18.00±0.06°C), the CM4-MG2's SST (17.42±0.10°C) is biased low by about 0.58°C (vs. 0.62°C low bias in CM4.0). In the AMIP simulations where the SST is prescribed (M. Zhao et al., 2018a), T_{air} is colder than the observation (Climatic Research Unit TS data-set version 4.01) by 0.62°C over the land (I. Harris et al., 2014), and colder than the ERA-Interim reanalyses (European Center for Medium Range Weather Forecasting Re-Analysis Interim) by 0.30°C over the ocean (Dee et al., 2011). Hence T_{air} in the CM4-MG2 piControl (12.78±0.14°C) is likely biased cold by 0.58°C or more (Figure 1 (c)). One reason for the temperature cold bias is the snow-covered glacial albedo that has been purposely tuned higher during the CM4.0 development, in order to encourage the formation of Antarctic bottom water (Held et al., 2019).

3.2 Historical Experiments

3.2.1 Atmosphere Climatology

We evaluate the atmosphere climatology over the period of 1980–2014. Three ensemble members of CM4-MG2 historical experiment are examined and compared to the CM4.0 counterpart experiments.

Figure 2(a) shows the global map of annual mean net downward shortwave flux or shortwave absorption (SWABS) at TOA from three CM4-MG2 historical ensemble mean. The observational reference is the Clouds and the Earth's Radiant Energy System-Energy Balanced and Filled climatology Edition 4.1 (CERES-EBAF-Ed4.1) shown in Figure 2(b) (Loeb et al., 2009, 2018). Model bias patterns are qualitatively similar between CM4-MG2 and CM4.0 (Figures 2(c)(d)). Negative biases are seen in the sub-Saharan Africa, western Indian Ocean, western Pacific storm track regions, tropical Atlantic, and near the Arctic (north of ∼60°N). Positive biases occur in the Southern Ocean (south of 60°S) and equatorial Pacific, and along the west coasts of South America, Africa, and North America, suggesting a lack of cloudiness. The lack of subtropical stratocumulus clouds off the west coasts have been a long-standing problem in the GFDL GCMs (Donner et al., 2011; M. Zhao et al., 2018a; Held et al., 2019; Dunne, Horowitz, et al., 2020).

This problem of coastal stratocumulus has improved noticeably with the introduction of the MG2 cloud microphysics in the AMIP mode simulations (Guo et al., 2021). More importantly, the fully coupled CM4-MG2 simulations successfully maintain this improvement (Figure 2(c)). Over three representative stratocumulus regions near Peru [80-90W, 10-20S], Namibia [0-10E, 10-20S], and California [120-130W, 20-30N] (Klein & Hartmann, 1993), the annual mean SWABS biases are 5.88, 10.72, and 7.30 W m^{-2} in CM4-MG2, about 10 W m^{-2} smaller than those in CM4.0 (16.10, 18.30, and 17.44 W m^{-2}). This indicates that the enhanced subtropical stratocumulus is a robust feature when MG2 is active. This enhancement is likely due to the Seifert and Beheng (2001) autoconversion scheme and the prognostic precipitation treatment, which suppress the autoconversion of cloud water to rain at low liquid water paths and help sustain the subtropical stratocumulus. The improvement in the shortwave off the west coasts is not only a significant regional improvement, but also has important implications especially for coupled simulations. It could reduce the warm SST biases of the underlying ocean and potentially alleviate the double ITCZ bias (Large & Danabasoglu, 2006), which will be discussed later.

An analogous figure for outgoing longwave radiation (OLR) is provided in Figure 3. The OLR spatial pattern at TOA from CM4-MG2 closely resembles the CERES-EBAF-

Ed4.1 observation. Comparison of two model biases gives an overall improvement in root-mean-square-error (RMSE, 5.75 vs. 6.31 W m^{-2} in CM4.0), and similar global mean bias (-2.19 vs. -2.37 W m^{-2} in CM4.0). Regionally, the OLR biases are closely correlated with convective precipitation biases in the tropics (M. Zhao et al., 2018a). Excessive OLR are present over the equatorial Pacific and Atlantic, and Amazon where dry biases are seen. Insufficient OLR occurs over West Indian Ocean, the Maritime continent, the tropical Pacific, and tropical South Atlantic, where wet biases are significant. As shown in Figures 3 and 4, the OLR biases appear to be larger where the precipitation biases are stronger, and vice versa. In the extratropics, the OLR tends to be underestimated in both models, but the underestimate is amplified in CM4-MG2 especially over the Arctic.

Both SWABS and OLR biases suggest too much cloudiness over the Arctic in CM4-MG2 (Figures 2 and 3). One reason is that the atmosphere is more humid (Figures 7(c)(d)), which favors more cloudiness since the large-scale cloud cover is parameterized as a function of relative humidity (Tiedtke, 1993). Table 2 provides the global annual means and RMSEs of clear-sky SWABS (SWABS_clr) and OLR (OLR_clr) at TOA, as well as shortwave and longwave radiative effect (SWCRE, LWCRE). CM4-MG2 shows lower OLR_clr and LWCRE than CM4.0 by about 1 W m^{-2}. Given higher water vapor content in CM4-MG2 (Figure 6), OLR_clr is effectively from the emissions at higher altitude (or colder temperature), and therefore lower.

Annual precipitation is exhibited in Figure 4 compared to Global Precipitation Climatology Project (GPCP) V2.3 (Adler et al., 2003, 2016). The global mean precipitation rates from both models are higher than the GPCP reference by $\sim$6-7%. Note that there exist systematic underestimations in the satellite retrievals, and the GPCP estimate may be biased low by $\sim$10% (Wild et al., 2013). As confirmed by Stephens et al. (2012), the GPCP precipitation is probably underestimated by $\sim$10% over tropical oceans, and by a larger percentage over mid-latitude oceans. The global mean precipitation rates of CM4-MG2 and CM4.0 (2.85 vs. 2.89 mm day^{-1}) differ by about 0.04 mm day^{-1}, similar to what is reported in the AMIP simulations (Guo et al., 2021). The lower precipitation rate in CM4-MG2 is likely associated with precipitation efficiency, defined as the ratio of surface precipitation rate to the sum of column-integrated vapor condensation and deposition rates (Sui et al., 2005, 2007). As discussed in Guo et al. (2021), the MG2 cloud microphysics shows lower precipitation efficiency than the RK microphysics. Because of the less efficient depletion of water vapor by precipitation, more vapor is present in CM4-MG2 (Figure 6).

Regional precipitation biases comprise dry Amazon and equatorial Pacific, wet tropical Africa and West Indian Ocean and tropical Pacific, as well as biases typically developed in the coupled simulations: wet maritime continent and double ITCZ. Double ITCZ is a common bias persisting in a number of state-of-the-art fully coupled GCMs (Held et al., 2019; Golaz et al., 2019; Voldoire et al., 2019; Kelley et al., 2020; Dunne, Horowitz, et al., 2020). It is manifested as a zonal band of excessive precipitation across the Southern Hemisphere tropics at about 8°S. One notable achievement of CM4.0 is the reduced double ITCZ bias compared to GFDL's previous-generation GCMs (Held et al., 2019). It is encouraging that CM4-MG2 further reduces it, with smaller wet biases (less reddish) in the Indian Ocean, the South Pacific Convergence Zone, and the tropical Atlantic Inter-Tropical Convergence Zone (Figures 4(c)(d)). To highlight the ITCZ improvement, we compare the zonal mean precipitation over the eastern Pacific (150W-90W). As shown in Figure 5, three CM4-MG2 ensemble members are closer to the GPCP observation over 2°S-10°S than the CM4.0 counterparts, indicating the reduced wet biases by this measure. This improvement is partly related to the enhanced subtropical stratocumulus clouds (Figure 2). The enhanced stratocumulus reflects more shortwave radiation back to space, and cools the underlying sea surface along the eastern boundaries of the subtropical ocean basins. The cooler SST might suppress local convection and result in less rainfall.

Figure 7 illustrates the model biases of surface air temperature and relative humidity at 2 m, and surface zonal wind at 10 m compared to ERA-Interim (Dee et al., 2011). The simulated surface air temperature appears to be biased cold, with a global mean value lower than the reanalysis by about 1.11 K. The cold bias partially arises from the boost of snow-covered glacial albedo, alleviating (or delaying) unrealistic superpolynya behavior in the Southern Ocean (Held et al., 2019). Albeit the cold bias is prevalent, the warm bias is present along the eastern boundaries of the subtropical ocean basins, as well as in the Ross Sea and Weddell Sea. Relative to CM4.0, CM4-MG2 shows marked improvements along the eastern boundaries largely due to the enhanced subtropical stratocumulus, but moderate degradation in the Southern Ocean. Comparison of 2 m relative humidity reveals positive biases in both models, especially in the high latitudes (Figures 7(c)(d)). CM4-MG2 shows larger biases there. One reason is possibly due to the less efficient precipitation formation and thus more humid atmosphere when MG2 is effective, as supported by higher water vapor path in CM4-MG2 (22.96 vs. 22.06 g m^{-2} in CM4.0) (Figure 6). The surface zonal wind biases at 10 m in CM4-MG2 and CM4.0 exhibit similar geographical patterns: positive biases over the Antarctic and negative biases over the Indian Ocean and equatorial Pacific (Figures 7(e)(f)). CM4-MG2 shows slightly larger global mean bias (-0.12 vs. -0.09 m s^{-1} in CM4.0), but slightly smaller RMSE (0.63 vs. 0.67 m s^{-1} in CM4.0). Additionally, Table 2 presents the global biases of surface wind stress (tau_x, tau_y), surface latent and sensible heat fluxes (LH_flx, SH_flx), and sea level pressure in the Northern and Southern Hemispheres (SLP_NH, SLP_SH). Both CM4-MG2 and CM4.0 show comparable global means and RMSEs, and are close to the ERA-Interim reanalyses.

We further examine the vertical profile of annually and zonally averaged zonal wind (Figure 8). Both CM4-MG2 and CM4.0 show the shift of midlatitude westerlies toward the equator, which is a common deficiency developed in the coupled GCMs (M. Zhao et al., 2018a; Held et al., 2019; Golaz et al., 2019). Both models underestimate the westerly throughout the midlatitude troposphere, and overestimate the trade winds in the tropics. Nevertheless the underestimate of the westerly and the overestimate of of trade winds are alleviated to some degree in CM4-MG2, leading to overall smaller RMSE (Figures 8(c)(d)). The corresponding temperature profile is shown in Figure 9. Both models share cold biases throughout the troposphere, consistent with the colder-than-observed SST discussed in Section 3.2.2. The cold bias is reduced in CM4-MG2, by about a factor of 2 in mid-upper troposphere over the tropical and mid-latitude regions, which is perhaps associated with more water vapor there (Figure 6). The positive biases are present in the stratosphere in both models, but reduced noticeably in CM4-MG2. Consequently, the RMSE of the CM4-MG2's temperature profile is smaller than that of CM4.0 (Figures 9(c)(d)).

3.2.2 Ocean and Sea Ice Climatology

The SST biases, relative to the Hadley Centre Sea Ice and Sea Surface Temperature (HadISST) data set over 1980-2014, averaged over three historical ensemble members for both CM4-MG2 and CM4.0 are displayed in Figure 10. The global mean biases (-0.64 vs. -0.63 K) and RMSEs (0.96 vs. 0.97 K) are comparable between CM4-MG2 and CM4.0. The geographical patterns of SST biases, to a large extent, are similar to surface air temperature biases (Figures 7(a)(b)). There are prevailing cold biases in the subtropical highs and their poleward margins, with warm anomalies in the Northwest Atlantic Ocean and oceanic upwelling regions along the west coasts of Africa, North and South America. Due to the enhanced coastal stratocumulus clouds, the warm biases along the west coasts are (marginally) improved in CM4-MG2, which is one of the possible reasons for the reduced double ITCZ bias (Figures 4(c)(d)). But this improvement is less significant than what is found in the surface air temperature, suggesting that lack of subtropical stratocumulus clouds are only part of the reasons for the SST warm biases along the west coasts. Other factors, for example, insufficient ocean upwelling, are likely contributors, too.

The seasonal cycle of sea ice extent (SIE) is shown in Figure 11. Both CM4-MG2 and CM4.0 agree favorably with passive microwave satellite observations from the National Snow and Ice Data Center (NSIDC) (Cavalieri et al., 1996). Both models well represent the magnitude and timing of Pan-Arctic SIE, with maxima in March and minima in September. But they both tend to overestimate the SIE, and CM4-MG2 further amplifies it, especially during the boreal winter and summer. Since both models adopt the same SIS2 code and tunings (Adcroft et al., 2019), we speculate that the overestimate amplification might be related to ice nucleation and cloud microphysical parameterizations, and more humid atmosphere in CM4-MG2, and thus excessive clouds over the Arctic as well as their interactions with the coupled system (Figures 2, 3, and 6). The overestimate of the Arctic sea ice is expected to enhance sea ice feedback, leading to higher climate sensitivity, which will be discussed in Section 3.5. Both CM4-MG2 and CM4.0 magnify the seasonal cycle of Pan-Antarctic SIE, with positive biases in the austral winter and negative biases in the austral summer (Figure 11(b)). These Pan-Antarctic SIE biases are also present in the GFDL SPEAR (Seamless System for Prediction and EArth System Research) simulations, and are suspected to be associated with too much shortwave absorption in summer or a missing subgrid ice ridging (dynamical thickening) scheme in SIS2 (Adcroft et al., 2019; Delworth et al., 2020).

3.3 Climate Variability

The evaluations so far have been mainly focused on the mean model climate. In this section, we will assess the model performance from the climate variability perspective.

3.3.1 Madden–Julian Oscillation (MJO)

The Madden-Julian Oscillation (MJO) is a key climate variability mode (Madden & Julian, 1971, 1972). It is the largest component of the intraseasonal (30- to 90-day) variability and a key feature of moist convection in the tropical atmosphere. Recent studies have shown that lower tropospheric moisture and its advection play a key role for the propagation and magnitude of the MJO (Benedict et al., 2014; Pritchard & Bretherton, 2014; Adames & Wallace, 2015; X. Jiang, 2017; H.-M. Kim, 2017; X. Jiang et al., 2020). The atmospheric moisture has increased remarkably in CM4-MG2 (Figure 6). More moisture may favor the development of convection due to the moisture-convection feedback (X. Jiang et al., 2020). Therefore it is expected that the MJO simulation will be impacted, even though the convection parameterization has not changed from CM4.0 to CM4-MG2.

Figure 12 shows the tropical symmetric power spectrum of OLR from 15°S to 15°N to assess the magnitude of MJO (Wheeler & Kiladis, 1999). The color shading regions indicate that the spectral power associated with MJO, Kelvin and other convective waves are greater than or equal to 1.2, which is significantly above the background noises. CM4-MG2 shows clearly stronger tropical wave activity. For example, in zonal wave number 1-3 (or frequency$\sim$0.025 day^{-1}), there is enhanced MJO. CM4-MG2 shows stronger eastward propagating OLR signals than CM4.0, and thereby agrees better with the AVHRR (Advanced Very High Resolution Radiometer) observation (Liebmann & Smith, 1996).

Further analyses are conducted by evaluating the life cycle of MJO. Figure 13 displays the composites of 20-100 day band-pass filtered daily anomalies in OLR and wind vector at 850 hPa (u850, v850) during the boreal winter season (November to April). The composites clearly illustrate the eastward propagation of convective signals, represented by the OLR anomalies. The negative OLR anomalies (associated with MJO) first develop over the Indian Ocean, get strengthened and pass through the Maritime Continent, then gradually decay and continue into the western Pacific. Both CM4-MG2 and CM4.0 well represent the traveling pattern of the MJO, compared to the ERA5 reanalysis. During the MJO life cycle, CM4-MG2 exhibits a larger magnitude of the OLR anomalies and/or stronger convective signals than CM4.0, and shows notable improvements in simulating the eastward propagation of the MJO.

3.3.2 El Niño-Southern Oscillation (ENSO)

The El Niño-Southern Oscillation (ENSO) is Earth's strongest year-to-year climate fluctuation, involving SST variations in the tropical Pacific that have major impacts on the global climate system (McPhaden, M J and A Santoso and W Cai, 2020). Thus it is critical for climate models to simulate realistic ENSO variability.

We conducted wavelet analyses (Torrence & Compo, 1998) for SST averaged over the Niño-3 region (150°W–90°W, 5°S–5°N), comparing the power spectra from observational reconstructions against those from the CM4-MG2 and CM4.0 piControl and historical ensemble (Figure 14). The observed spectrum, based on the NOAA Extended Reconstructed Sea Surface Temperature, version 5 (ERSSTv5) observational reanalysis (Huang et al., 2017), shows a strong annual peak and a broad interannual peak spanning 2–8 years (Larkin & Harrison, 2002; Kessler, 2002; Wittenberg, 2009; Wittenberg et al., 2014). For CM4.0, the simulated spectra closely resemble the observations, with a broad interannual peak. For CM4-MG2, the spectra show a stronger ENSO with a somewhat longer period than observed. The ENSO period peaks near 3.5-4.0 years for CM4-MG2, while it is 3.3 years for observations and CM4.0. In both CM4.0 and CM4-MG2, the simulated historical annual cycle of Niño-3 SST is slightly stronger than observed; and moving from pre-industrial to historical forcings, in both models the ENSO strengthens while the annual cycle weakens. Given the excellent spectra in the CM4.0 historical simulations, it is somewhat disappointing that the enhanced subtropical stratocumulus in CM4-MG2 results in an apparent overestimate of the ENSO amplitude. Yet given the numerous competing coupled feedbacks involved in ENSO, it is often the case that improvements in one model component can unmask shortcomings in other components (Wittenberg et al., 2018; Guilyardi et al., 2020). These shortcomings will need to be identified and addressed via additional iterations of coupled model development.

3.3.3 Atlantic Meridional Overturning Circulation (AMOC)

Figure 15 shows the maximum Atlantic Meridional Overturning Circulation (AMOC) at 26°N, which was estimated by integrating volume transport down from the ocean surface. The mean AMOC strengths, from the CM4-MG2 and CM4.0 historical ensemble members over the period of 2004–2014, are about 16.38 Sv and 15.82 Sv, which are close to the direct observation from the RAPID array ($\sim$16.9$\pm$3.35 Sv) (Moat et al., 2020). In the historical simulations, the modelled AMOC exhibits a strengthening trend from 1940 to 1980, but after peaking around 1980, it shows a weakening trend (Figure 15(a)). These trends are generally consistent with the simulated AMOC variations in the state-of-the-art GCMs from CMIP6 (Hassan et al., 2021; Menary et al., 2020), and are likely to be related to the compensating effects between aerosols and greenhouse gases (GHGs) (Delworth & Dixon, 2006; Hassan et al., 2021; Menary et al., 2020). Increasing GHGs contributes to the weakening of the AMOC, while aerosols impose opposite effects and offset the GHG-induced weakening. The build-up of anthropogenic aerosols increases the strength of AMOC prior to 1980, and the following AMOC weakening is likely due to the reduced aerosol emissions and increasing GHGs.

Figure 15(b) provides the time series of the 10-yr running average AMOC from the piControl simulations. The mean AMOC strengths from both models are comparable to the observed mean, with slightly stronger AMOC in CM4-MG2 than CM4.0 (17.26 vs. 16.71 Sv). But the multidecadal variability of the modelled AMOC is underestimated, evidenced by lower standard deviations of 0.54 Sv for CM4-MG2 and 0.60 Sv for CM4.0 versus 1.37 Sv for indirectly inferred observations (Yan et al., 2018). Furthermore, the simulated forced multidecadal AMOC variations (Figure 15(a)) are opposite to the historical multidecadal AMOC variations inferred from the observed AMOC fingerprints (i.e. a negative phase during 1970s and 1980s and a positive phase during 1960s and post-1990), which are more likely dominated by internal variability (Yan et al., 2019). This discrepancy with the observational records is consistent with the muted internal multidecadal AMOC variability in this model. The lower multidecadal variability might be associated with the buoyancy forcing (W. M. Kim et al., 2017), and is also suspected to be partially related to the wind forcing (J. Zhao & Johns, 2014; Yan et al., 2018). Recent reconstructions of the long-term mean AMOC structure suggests that the Arctic is the northern terminus of the mean AMOC (Zhang & Thomas, 2021), and the simulated lower multidecadal AMOC variability is likely related to the underestimated multidecadal Arctic salinity variations in climate models due to the model biases in the Arctic (Rosenblum et al., 2021). Nevertheless, detailed discussion on the underlying reasons for the muted multidecadal AMOC variability is beyond the scope of this study.

3.4 Temperature Evolution and Aerosol Radiative Forcing

Figure 16(a) provides the time evolution (1850-2014) of the global mean surface temperature anomaly from the CM4-MG2 and CM4.0 historical ensembles, as well as the comparison against the observational estimate: NASA Goddard Institute for Space Studies Surface Temperature product version 4 (GISTEMP v4) (GISTEMP-Team, 2019; Lenssen et al., 2019). The temperature anomaly is the 5-year running average relative to the 1880–1900 period, which is the first 20-yr of the GISTEMP data. The blue curve is the CM4-MG2 three historical ensemble mean, and the shaded region is the ensemble range. The red curve is the CM4.0 three historical ensemble mean. The black curve is the observational estimate. The letters above the x-axis indicate major volcano events. Each event results in a dip in temperature. From 1880 to 2014, the overall bulk global warming from both models agrees well with observations, although warmer than the observation before 1940 and colder after 1960. The cold bias persists until 2010 when it is virtually cancelled out by the abrupt warming starting around 1990.

More details about the warming are displayed in the difference of temperature anomaly between the Northern Hemisphere (NH) and the Southern Hemisphere (SH) (Figure 16(b)). From 1920 to about 1980, the NH exhibits stronger warming than the SH from the GISTEMP observation, but neither CM4-MG2 nor CM4.0 captures this hemispheric warming asymmetry, suggesting insufficient modelled warming (or too strong cooling) in the NH. After 1980, both models (especially CM4.0) show a rapid warming trend in the NH, similar to the abrupt warming in the global mean temperature since 1990 (Figure 16(a)). The rapid warming trend might be related to aerosol radiative effect and climate sensitivity.

The time series of aerosol radiative flux perturbation (RFPs) for the NH and the SH are shown in Figure 17. The RFP is estimated as the change in the TOA net radiation from a pair of climatological simulations with identical SST and sea ice but different (present-day or pre-industrial) radiative forcing agents and their precursors (Lohmann et al., 2010; Golaz et al., 2011; Hansen et al., 2014; Forster et al., 2016). As anthropogenic aerosol emissions increase remarkably from 1920 to 1990, the aerosol RFP gets stronger (more negative) especially in the NH. During the period of 1970-1990, the RFP in the NH reaches -1.41 W m^{-2} and -1.65 W m^{-2} for CM4-MG2 and CM4.0, respectively. Such strong aerosol cooling is capable of offsetting or partially offsetting the greenhouse warming. After 1990, the aerosol RFP declines quickly (or becomes less negative) especially for CM4.0. The quick decline in the aerosol cooling, along with the rapid increase in the greenhouse warming, leads to an abrupt warming trend as shown in Figure 16. In addition to the aerosol radiative effect, climate sensitivity is another important factor influencing the global warming, which will be discussed in Section 3.5.

3.5 Climate Sensitivity and Cloud Feedback

Two idealized CO_2 forcing simulations: CO_2 concentration increasing 1% per year (1pctCO2) and abruptly quadrupled CO_2 (abrupt-4XCO2) (see Table 1), were conducted to evaluate climate sensitivity. Climate sensitivity is an important metric to understand the trajectory of the 20th century warming (Figure 16), as well as the climate projection and long-term climate outcomes of the 21th century and beyond. A model with a higher climate sensitivity is likely to yield a larger temperature change for given anthropogenic forcing.

Transient Climate Response (TCR) is a primary measure of climate sensitivity under increasing CO_2 scenario, referring to the warming at the time of CO_2 doubling (around Year 70) in the 1pctCO2 experiment (Table 3). Figure 18 (a) illustrates the time evolution of global annual mean surface air temperature change (ΔT). In response to increasing CO_2 concentration, CM4-MG2 warms less than CM4.0. The TCR, from the difference of 20-year averages (i.e., Year 61-80) between the 1pctCO2 and piControl, is about 10% lower in CM4-MG2 than that in CM4.0 (1.85 vs. 2.05 K). In Year 140 when CO_2 is quadrupled, the warming is about 4.16 K in CM4-MG2 but reaches 5.10 K in CM4.0, although both well above twice their corresponding TCRs. Furthermore, CM4-MG2 exhibits weaker warming than CM4.0 in the abrupt-4xCO2 experiment (Figure 18 (b)), echoing the less warming shown in the 1pctCO2 experiments.

Another benchmark sensitivity metrics is equilibrium climate sensitivity (ECS), defined as the equilibrium global surface temperature change in response to CO_2 doubling. But the evaluation of ECS is usually expensive computationally, because it takes thousands of model years for a coupled GCM to achieve equilibrium or steady state. As shown in Figure 18 (b), the 150-year simulation of the abrupt-4xCO2 is far from equilibrium. Nevertheless, Winton et al. (2020) extended the abrupt-4xCO2 experiment to 300 years and yielded an estimate of ECS of about 5.0 K for CM4.0. Following Dunne, Winton, et al. (2020), we estimated the ECS of 4.52 K and 4.89 K for CM4-MG2 and CM4.0, respectively. Another comparable and widely used alternative is effective climate sensitivity (EffCS), following the method of J. Gregory et al. (2004). This method is to simply regress the top of atmosphere net radiative flux change (ΔN) against ΔT. From the linear regression for all 150 years of the abrupt-4xCO2 experiment, EffCS can be diagnosed as the half of the ΔT-axis intercept (i.e., half of x-axis intercept in Figure 18 (c)). The half is to evaluate EffCS with respect to a CO_2 doubling according to its definition. With this method, the estimates of EffCS are 3.31 K and 3.91 K in CM4-MG2 and CM4.0, respectively. CM4-MG2's lower EffCS is consistent with its lower TCR. The introduction of the MG2 cloud microphysics appears to reduce climate sensitivity.

In order to understand why the climate sensitivity is reduced, we diagnose effective radiative forcing from a doubling of CO_2 (EffF_{2x}) and climate feedback parameter (λ_{net}) under the assumption of EffCS=-Eff$F_{2x}/\lambda_{\mathrm{net}}$. Again Eff$F_{\mathrm{2x}}$ and λ_{net} are derived by regression, and calculated as the half of the ΔN-axis intercept and the slope of the linear regression line (Figure 18). It is not surprising that lower EffCS in CM4-MG2 results from weaker EffF_{2x}, and more importantly from smaller (more negative) λ_{net} (Table 3). This is similar to what is reported for the CMIP6 GCMs (as compared to the earlier CMIP5 generation GCMs). The combination of feedback and forcing results in higher EffCS in CMIP6: higher (less negative) feedback accounts for 60% increase of EffCS while stronger forcing only contributes to 20% increase (Zelinka et al., 2020). Hence, λ_{net} is a major contributor to the change in EffCS.

The global map of λ_{net} is displayed in Figures 19(b)(c). The spatial patterns of λ_{net} are similar for CM4-MG2 and CM4.0. λ_{net} is mostly negative, and becomes positive in the North Asia, Northern Canada, tropical East Pacific, and Southern Ocean. The zonally averaged λ_{net} in CM4-MG2 is generally smaller (more negative) than that in CM4.0, except for northern subpolar where λ_{net} peaks (Figure 19(a)). The larger λ_{net} around 70°N in CM4-MG2 is mainly because of shortwave clear-sky feedback (λ_{SWclr}), after decomposing λ_{net} into longwave and shortwave clear-sky (λ_{LWclr}, λ_{SWclr}), and cloud radiative effect (λ_{CRE}) components. As shown in Figures 19(d)(e)(f), λ_{SWclr} ranges from neutral to strongly positive. CM4-MG2 exhibits larger λ_{SWclr}, especially poleward of 60°N. We attribute the larger λ_{SWclr} mostly to the decrease of surface albedo due to changes in snow cover and sea ice extent in the Arctic with warming. Both models overestimate the Arctic sea ice extent, but CM4-MG2 amplifies the overestimate (Figure 11(a)). This amplification further enhances the positive sea ice albedo feedback, and therefore increases the feedback in the Arctic. Note that the longwave clear-sky feedback (λ_{LWclr}) does not differ much between CM4-MG2 and CM4.0 (Table 3). So we will not discuss it further. The difference in λ_{net} largely stems from the differences in λ_{SWclr} and λ_{CRE}.

The positive λ_{SWclr} in the high-latitudes is partly balanced by the cloud radiative effect feedback (λ_{CRE}) (Figures 19(g)(h)(i)). Both CM4-MG2 and CM4.0 show strong negative λ_{CRE} in the Arctic and Southern Ocean, counteracting the strong positive λ_{SWclr}. Although both models share similar spatial patterns, for example, noticeably bimodal distribution (i.e., negative peaks at poleward of about 60°S and 70°N), CM4-MG2 overall exhibits weaker or more negative λ_{CRE} (with the global mean decreasing from 0.18 to -0.02 W m^{-2} K^{-1}). Note that the differences in λ_{CRE} cannot be simply ascribed to the differences in clouds (or cloud feedback). Some changes in cloud radiative effect come from the cloud mask of clear sky fluxes, rather than from cloud changes. So λ_{CRE} does not truly represent cloud feedback (λ_{CLD}). In order to better account for cloud masking effects, we then estimate λ_{CLD} using the radiative kernels described in Soden et al. (2008), instead of the linear regression. These radiative kernels were estimated using a control integration of the GFDL AM2 (GFDL Global Atmosphere Model Development Team, 2004), whose radiation algorithm is consistent with what is adopted in CM4.0 and CM4-MG2. Compared to λ_{CRE}, λ_{CLD} is systematically more positive. Its global mean is enhanced by about 0.5 W m^{-2} K^{-1} (Table 3), similar to $\sim$0.3-0.4 W m^{-2} K^{-1} reported by Soden et al. (2004). The differences in λ_{CLD} between CM4-MG2 and CM4 mainly occur in the extratropics (e.g., poleward of 30°S) (Figures 20(a)(b)(c)), and the global mean λ_{CLD} is reduced in CM4-MG2 (0.49 vs. 0.66 W m^{-2} K^{-1} in CM4.0). It is noted that given the approximations of the kernel technique, there often exists a residual feedback term, which is the difference between λ_{net} and the sum of kernel-derived components (Table 3). The residual term here is acceptably small ($\sim$0.1-0.2 W m^{-2} K^{-1}). So cloud feedback results are not expected to change qualitatively.

In order to better understand the reduction in λ_{CLD}, we analyze low-level cloud amount and liquid water path (LWP) changes against ΔT (Zelinka et al., 2020). The low-level cloud amount tends to decrease (positive feedback) while the LWP tends to increase (negative feedback) as the climate warms. Figures 20(g)(h)(i) show the zonal average and geographic distribution of the LWP change. CM4-MG2 exhibits stronger LWP increase in the tropical west Pacific. We suspect it might be associated with less efficient ice nucleation and more liquid clouds with warming. CM4-MG2 seems to experience weaker LWP increase in the extratropics (especially in the Southern Ocean) (Figures 20(h)(i)). The weaker increase is probably related to higher liquid fraction (or more super-cooled water) in CM4-MG2 (Andrews et al., 2019; Zelinka et al., 2020). When MG2 and dust-dependent ice nucleation are active, the supercooled liquid fraction tends to be higher especially for the mixed-phase clouds of temperature between -30 °C and -10 °C (See Figure 11 in Guo et al. (2021)). The smaller LWP increase is supposed to reduce cooling, leading to weaker negative (or stronger positive) cloud feedback in CM4-MG2.

However, the LWP increase with warming is accompanied by low-level cloud amount decrease, consistent with what is reported in the AMIP mode simulations (M. Zhao et al., 2016). As shown in Figure 20(d), both CM4-MG2 and CM4.0 exhibit reduced cloud amount with warming. But the cloud amount reduction is smaller in CM4-MG2 (-1.62% K^{-1} vs. -2.04% K^{-1} in CM4.0), and thus less positive cloud feedback. The net cloud feedback turns out to be 0.49 W m^{-2} K^{-1} in CM4-MG2, lower than 0.66 W m^{-2} K^{-1} in CM4.0 (Table 3). The decrease in low-level cloud amount is suspected to be related to precipitation efficiency (M. Zhao et al., 2016). In order to explore the impacts of precipitation efficiency, we have conducted a pair of present-day simulation and global warming simulation with SST uniformly warmed by 2 K following Cess and Coauthors (1990), and compared precipitation efficiency changes in a warmer climate. The precipitation efficiency is calculated as the ratio of surface precipitation rate to the sum of column-integrated vapor condensation and deposition rates (Sui et al., 2005, 2007). It is found that clouds occur less frequently and precipitation efficiency decreases with warming. The precipitation efficiency is reduced by about 0.72% K^{-1} with MG2, and by about 0.49% K^{-1} with RK, respectively. The stronger reduction in precipitation efficiency with MG2 results in weaker decrease in the low cloud amount (see Figure 4 in M. Zhao et al. (2016)), which contributes to less warming (or less positive cloud feedback). This is further supported by smaller (more negative) Cess feedback when MG2 is active (-2.02 vs. -1.77 W m^{-2} K^{-1} in Table 3). Although recent studies showed that the Cess experiments provide useful insight on cloud feedback (Ringer & et al., 2006; Ringer et al., 2014; Brient et al., 2015), a caveat is that the Cess approach assumes uniform SST warming and ignores important feedbacks, such as sea ice feedback and polar amplification. Hence the Cess feedback might underestimate the feedback of high latitude processes. The impacts of precipitation efficiency on cloud feedback (or climate sensitivity) in the fully coupled mode need more research in the future.

4 Summary

This paper describes the model performance and simulation characteristics of a fully coupled atmosphere-ocean-land-sea ice model configuration: CM4-MG2, and comparisons to the base model: CM4.0. CM4-MG2 and CM4.0 share the same ocean, sea ice, and land components. They only differ in the atmospheric component, or more specifically cloud microphysics: two-moment Morrison-Gettelman bulk microphysics with prognostic precipitation (MG2) vs. one-moment^{+} Rotstayn-Klein bulk microphysics with diagnostic precipitation (RK), and the mineral dust and temperature-dependent ice nucleation scheme. Based on a suite of CMIP6 DECK and historical simulations, model mean climate, climate variability, the 20th century simulation, and climate sensitivity have been examined and evaluated against available observations and reanalyses.

The CM4-MG2 mean climate is close or better relative to CM4.0 in terms of RMSE metrics. For some fields (e.g., OLR, temperature profile), the global RMSE is lower in CM4-MG2. The achievements include enhanced subtropical stratocumulus and reduced double ITCZ bias. The enhancement is a robust feature in both atmosphere-only and coupled simulations when MG2 is active. This is likely attributed to more realistic prognostic precipitation treatment and autoconversion parameterization (Guo et al., 2021). The enhanced stratocumulus also ameliorates the underlying SST warm bias along the west coasts of continents, and helps reduce rainfall and double ITCZ bias (Large & Danabasoglu, 2006). The degradation is the overestimate of the Arctic sea ice extent.

The simulated climate variability generally compares favorably with observations. CM4-MG2 shows stronger eastward propagating MJO signals than CM4.0, and agrees better with observations and reanalyses. One plausible reason is that the atmosphere is more humid in CM4-MG2 due to lower precipitation efficiency of MG2. The improved MJO simulation is expected to benefit the sub-seasonal to seasonal prediction (Xiang et al., 2021). Compared to the credible ENSO simulation with CM4.0, CM4-MG2 overestimates the spectral power and period lengths of ENSO. The modelled mean AMOC strength is in good agreement with the direct observation of RAPID, although its variability is muted. Both CM4-MG2 and CM4.0 simulate a strengthening trend of AMOC from 1940 to 1980 and a compensating reduction thereafter, due to the compensating effects between aerosols and GHGs. However, these simulated forced multidecadal AMOC variations are opposite to those inferred from the observed AMOC fingerprints over the second half of the twentieth century, which show a negative phase during 1970s and 1980s and a positive phase during 1960s and post-1990 and are more likely dominated by internal variability (Yan et al., 2019). This discrepancy between CM4.0/CM4-MG2 and the observational records is consistent with the fact that CM4.0/CM4-MG2 has insufficient internal multidecadal AMOC variability.

Both CM4-MG2 and CM4.0 are capable of simulating the bulk warming of the 20th century. But the temporal evolution of historical warming, to some extent, departs from the observation: insufficient warming from 1960 to 1990 and too rapid warming from there on. An analysis on the hemispheric warming asymmetry between the NH and SH reveals the cold bias (or insufficient warming) in the NH prior to 1980 and subsequently abrupt warming, especially in CM4.0. The abrupt warming and warming asymmetry are also concerns for a number of CMIP6 GCMs (Golaz et al., 2019; Held et al., 2019; Danabasoglu et al., 2020; C. Wang et al., 2021). Likely reasons are associated with aerosol radiative forcing and climate sensitivity (C. Wang et al., 2021). CM4-MG2 exhibits weaker (less negative) aerosol forcing than CM4.0 particularly in the NH, because the prognostic precipitation treatment in MG2 suppresses the dependency of rain formation on cloud drop size or number concentration (Posselt & Lohmann, 2008, 2009; Gettelman et al., 2015b; Guo et al., 2021).

CM4-MG2 exhibits lower climate sensitivity than CM4.0. The transient climate response (TCR) is 1.85 K and 2.05 K for CM4-MG2 and CM4.0, respectively. The effective climate sensitivity (ECS) is 3.31 K and 3.91 K, which are well within the expert estimated range (2.3–4.7 K) (Sherwood et al., 2020). It is not surprising that lower sensitivity largely results from weaker cloud feedback (Webb et al., 2006; Andrews et al., 2012), especially shortwave component (Zelinka et al., 2020; C. Wang et al., 2021). We further analyzed the changes of LWP and low-level cloud amount, and found that when the climate warms, CM4-MG2 exhibits weaker LWP increase and weaker low cloud amount decrease than CM4.0, especially over the Southern Ocean. These changes are related to higher liquid fraction and stronger precipitation efficiency reduction with warming in CM4-MG2. As demonstrated by M. Zhao et al. (2016), precipitation efficiency could strongly affect the model estimate of Cess sensitivity in the AMIP mode. We suspect that the lower climate sensitivity in CM4-MG2 is also partly associated with precipitation efficiency. A more detailed investigation on the impacts of precipitation efficiency in the coupled mode is beyond the scope of current paper and warrants further research.

The MG2 cloud microphysics is more expensive computationally than the RK scheme, mainly due to additional prognostic tracers (e.g., number and mass of rain and snow, ice crystal number concentration) and substepping in cloud microphysics. As a result, the overall computational cost increases by about 10% in the AMIP mode simulations (Guo et al., 2021). However, in the fully coupled simulations, there are barely any noticeable slow-down because of the loading balance between different model components. In the current configuration of CM4.0, the wall clock time for the ocean/sea ice component is slower than that of atmosphere/land component by 10% or more. This probably masks the slowdown caused by the MG2 microphysics in the CM4-MG2 atmospheric component.

While the CM4-MG2 coupled global simulations are promising, there are areas for further improvements and/or exploration. The MG2 microphysics enhances the subtropical stratocumulus clouds, but there is still lack of stratocumulus especially along the coasts, as shown by noticeable positive biases in the shortwave absorption. Refined vertical resolution can better resolve sharp temperature and moisture gradients of inversion, and is expected to better represent subtropical boundary layer clouds (Bogenschutz et al., 2021; Lee et al., 2021). The trajectory of the 20th century warming and hemispheric warming asymmetry somewhat deviates from the observation. This could be related to aerosol effects, climate sensitivity, among others. Given that climate sensitivity in CMIP6 GCMs increases substantially and that high sensitivity likely degrades the quality of the 20th century simulation and future projection, further research on climate sensitivity or cloud feedback is a high priority. Meanwhile, a credible 20th century simulation under the temperature trend constraint does not necessarily satisfy the "bottom–up" process level constraint such as cloud droplet size and cloud water phase partition (Golaz et al., 2013; Suzuki et al., 2013; Bodas-Salcedo et al., 2019). Future model development also needs to take the observational constraints on process level into account, in addition to the "top–down" constraints such as TOA radiative fluxes, atmospheric state, and temperature trend (Held et al., 2019; Mülmenstädt et al., 2020).

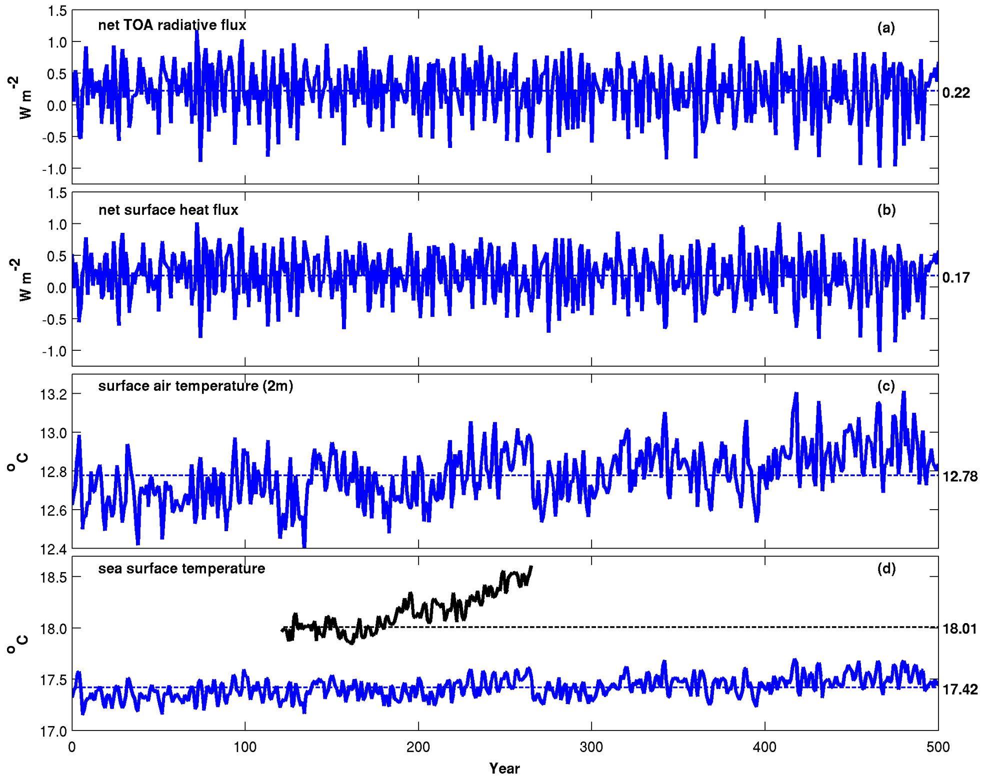

Figure 1. Time series of annual (a) global mean net radiative flux at top-of-atmosphere (TOA) (positive down), (b) global mean net heat flux at surface, (c) global mean surface air temperature, and (d) global mean sea surface temperature (SST). Blue solid lines represent the 500-yr time series of the CM4-MG2 piControl experiment. Blue dashed is the 500-yr average of the CM4-MG2 piControl. Black solid is the time evolution (1870-2014) of the Hadley Centre Sea Ice and Sea Surface Temperature data set (HadISST) (Rayner et al., 2003). Black dashed line is the time average of HadISST over 1880-1990.

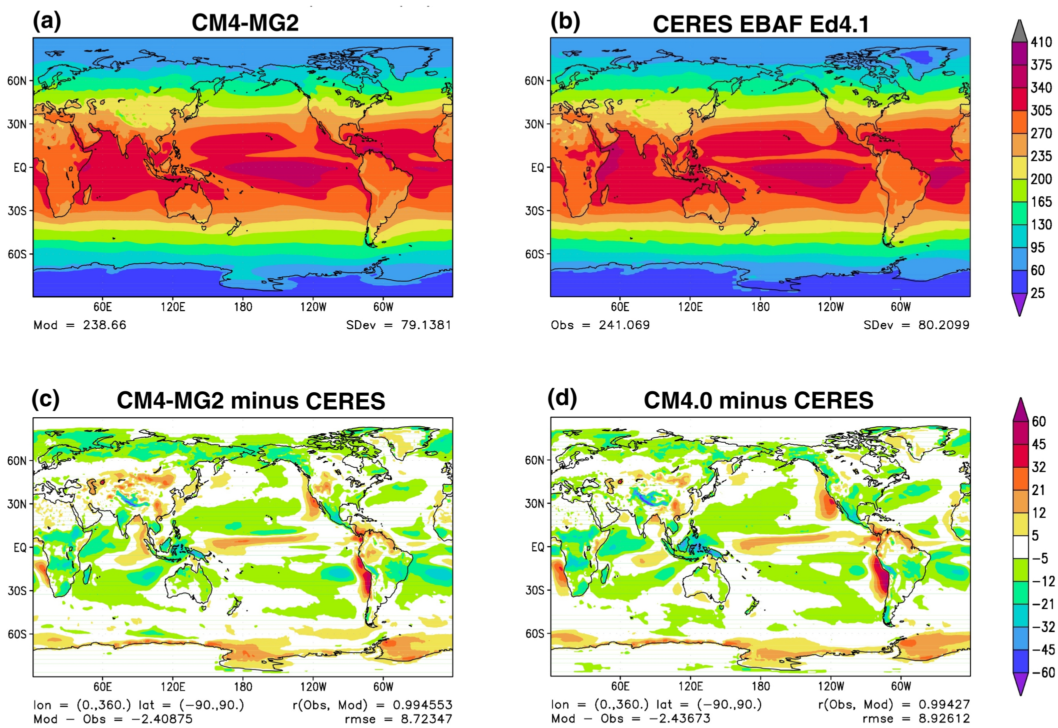

Figure 2. Annual-mean net downward shortwave flux or shortwave absorption (SWABS, W m^{-2}) at top-of-atmosphere (TOA) from (a) three-member ensemble mean of CM4-MG2 historical experiment for 1980-2014, (b) CERES–EBAF Ed4.1 averaged for 2000-2015, (c) CM4-MG2 model error (CM4-MG2 historical ensemble mean minus CERES), and (d) CM4.0 model error.

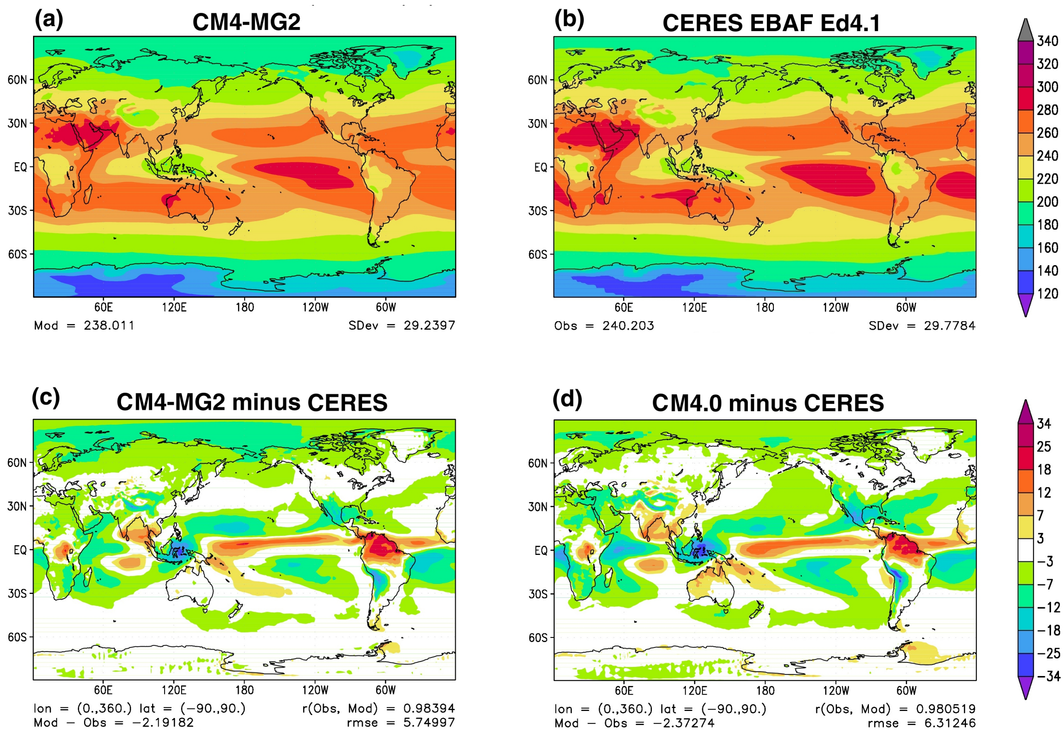

Figure 3. As in Fig. 2 but for outgoing longwave radiation (OLR, W m^{-2}) at top-of-atmosphere (TOA)

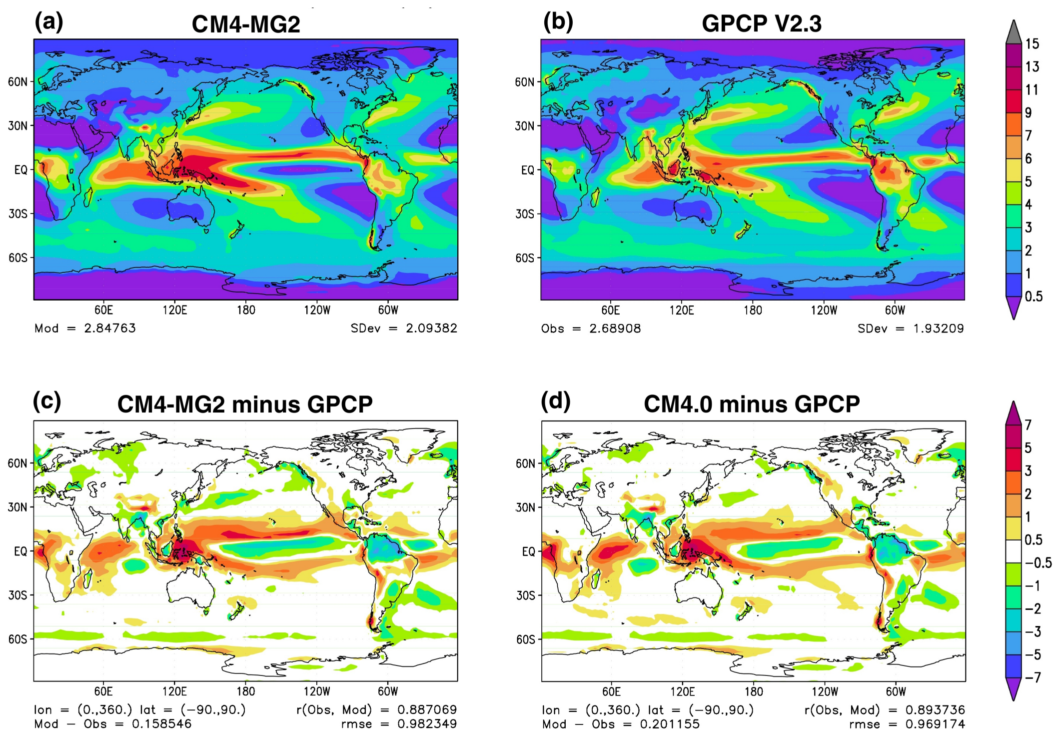

Figure 4. Annual-mean surface precipitation rate (mm day^{-1}) for (a) three-member ensemble mean of CM4-MG2 for 1980-2014, (b) GPCP v2.3 averaged for 1980-2015, (c) CM4-MG2 model error (CM4-MG2 historical ensemble mean minus GPCP), and (d) CM4.0 model error.

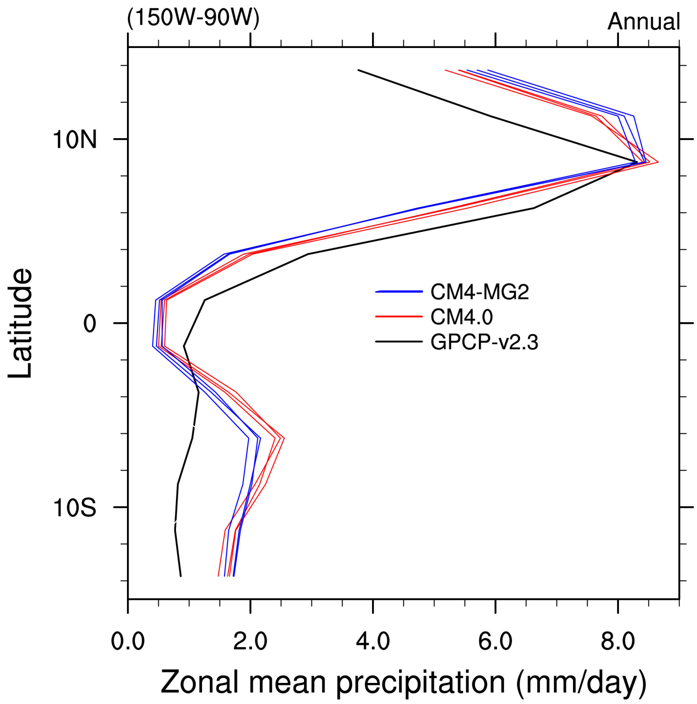

Figure 5. Annual zonal mean precipitation rate in the tropical Eastern Pacific averaged over longitudes 150W-90W as a function of latitude: GPCP v2.3 observations (in black), and three CM4-MG2 (in blue) and three CM4.0 (in red) ensemble members averaged over the years 1980-2014.

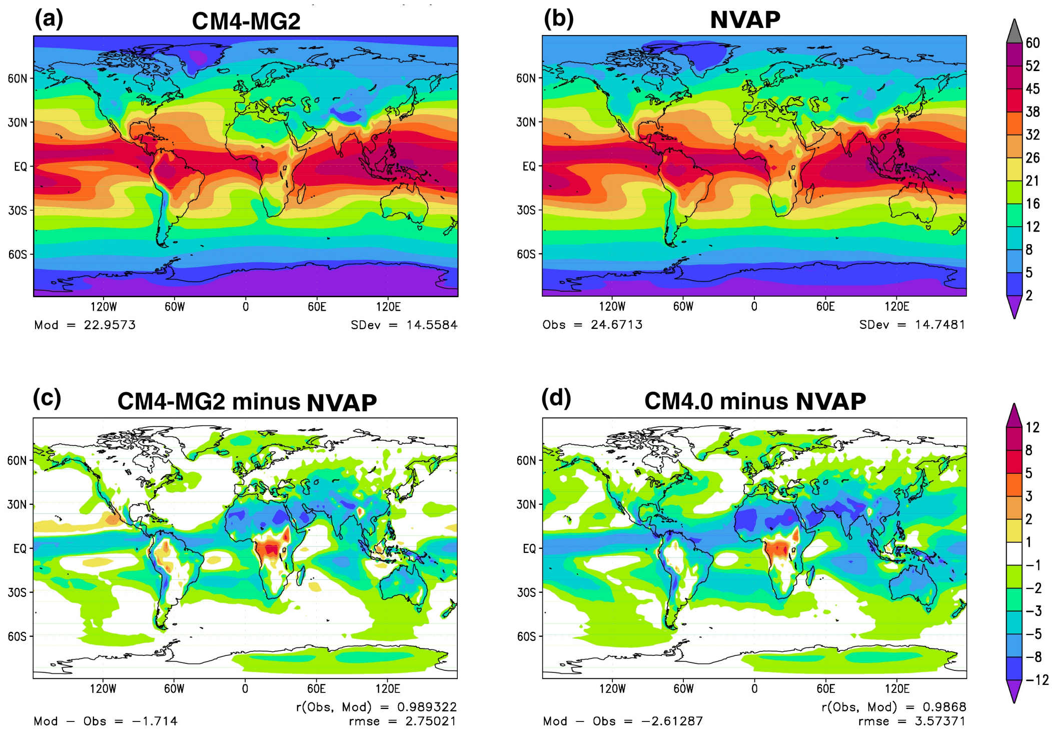

Figure 6. Annual-mean column-integrated water vapor path (WVP, kg m^{-2}) from (a) three-member ensemble mean of CM4-MG2 for 1980-2014, (b) the NASA Water Vapor Project (NVAP) total column water vapor data sets (Vonder Haar et al., 2012), (c) CM4-MG2 model error (CM4-MG2 historical ensemble mean minus NVAP), and (d) CM4.0 model error.

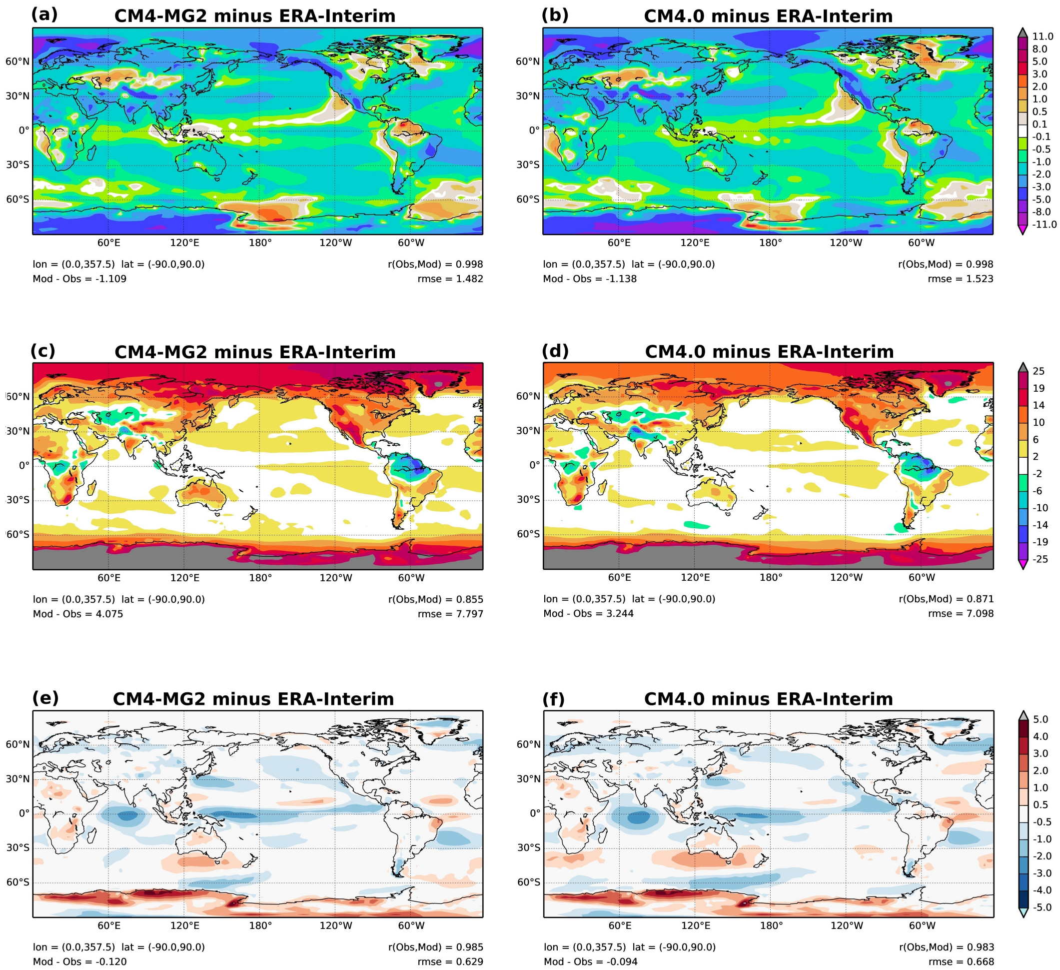

Figure 7. Model biases for 1980-2014 relative to ERA-Interim of near-surface (2 m) air temperature (°C) in (a) and (b), near-surface (2 m) relative humidity (%) in (c) and (d), and near-surface (10 m) eastward component of wind (m s^{-1}) in (e) and (f).

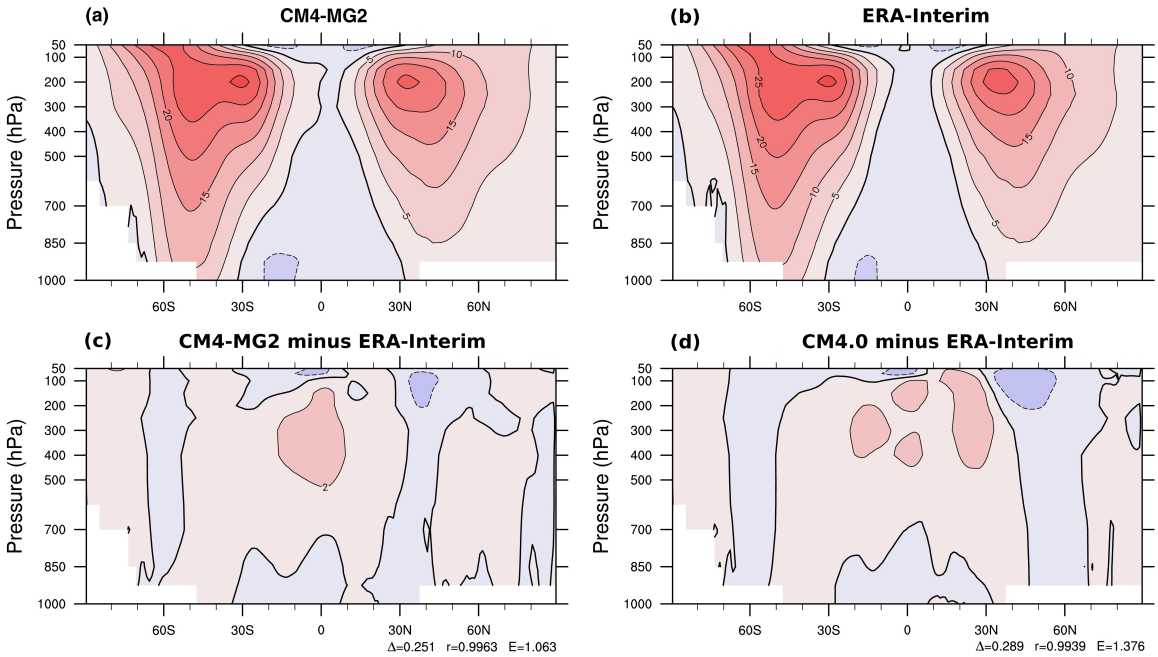

Figure 8. Annually and zonally averaged zonal wind (m s^{-1}) for (a) three-member ensemble mean of CM4-MG2 for 1980-2014, (b) ERA-Interim reanalysis averaged over 1980-2014, (c) CM4-MG2 model error (CM4-MG2 historical ensemble mean minus ERA-Interim), and (d) CM4.0 model error.

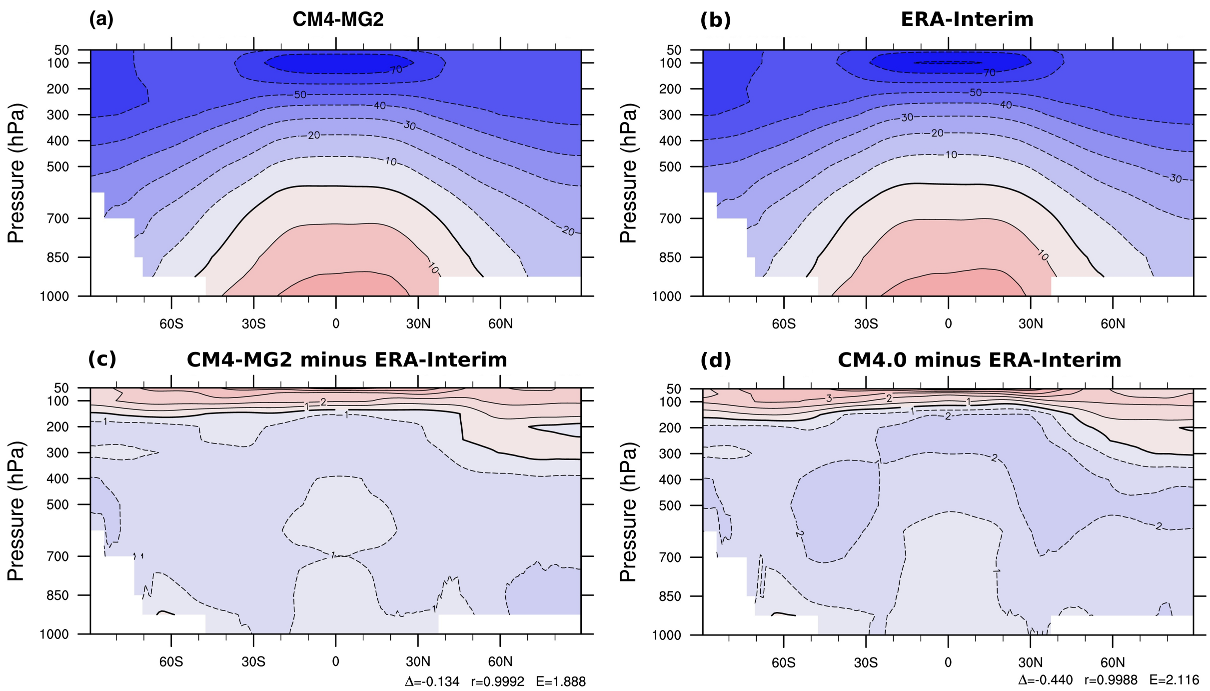

Figure 9. As in Fig. 8 but for annually zonally averaged temperature (°C).

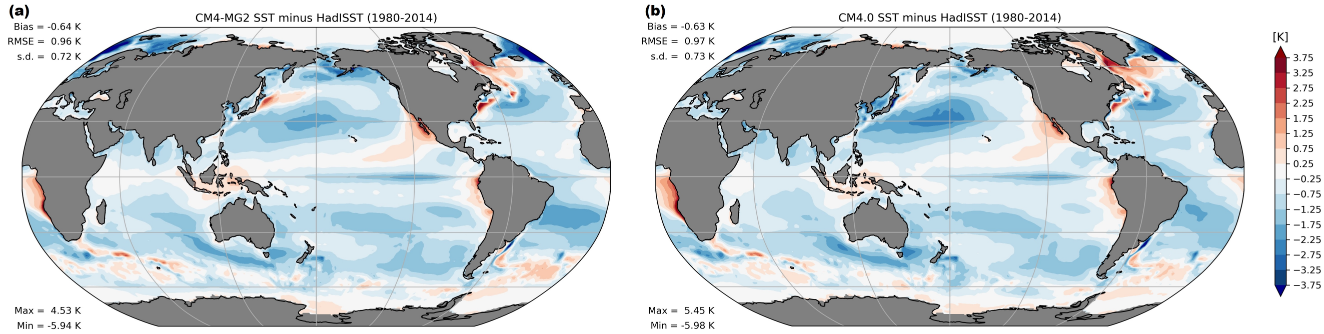

Figure 10. Sea surface temperature biases (K) in CM4-MG2 (a) and CM4.0 (b), from three historical ensemble members for 1980-2014, relative to the HadISST data set (Rayner et al., 2003) for the same time period.

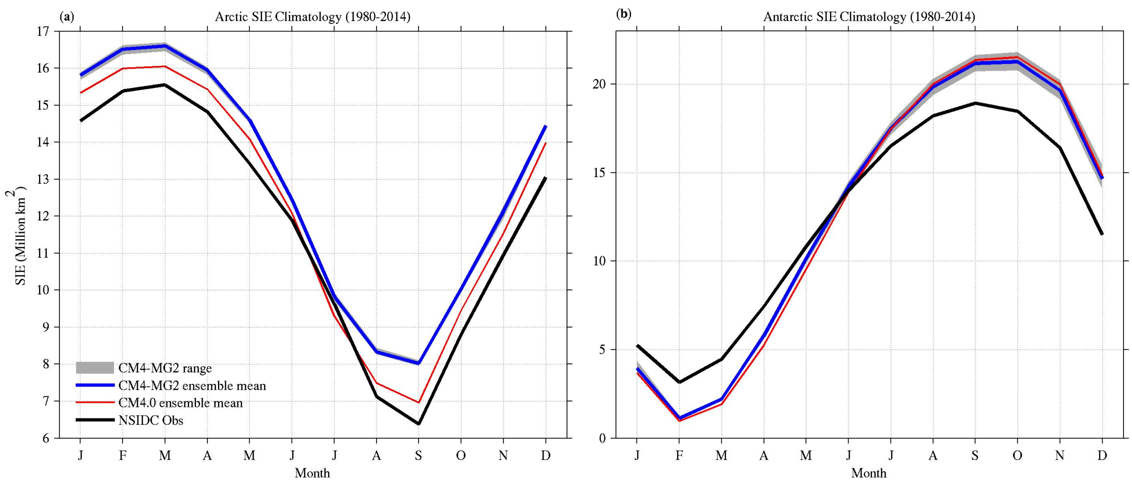

Figure 11. Sea ice extent (SIE) monthly climatologies (million km^2) for Pan-Arctic in (a) and Pan-Antarctic in (b) from three CM4-MG2 historical ensemble mean (thick blue), the spread based on the minimum and maximum values of three CM4-MG2 ensemble members (gray shaded), three CM4-MG2 historical ensemble mean (thin red), and satellite observations (black) from the National Snow and Ice Data Center (NSIDC) (Cavalieri et al., 1996). Pan-Arctic and Pan-Antarctic SIE are defined as the areal sum of all grid points whose sea ice concentration (SIC) exceeds 15% in the Northern and Southern Hemispheres, respectively.

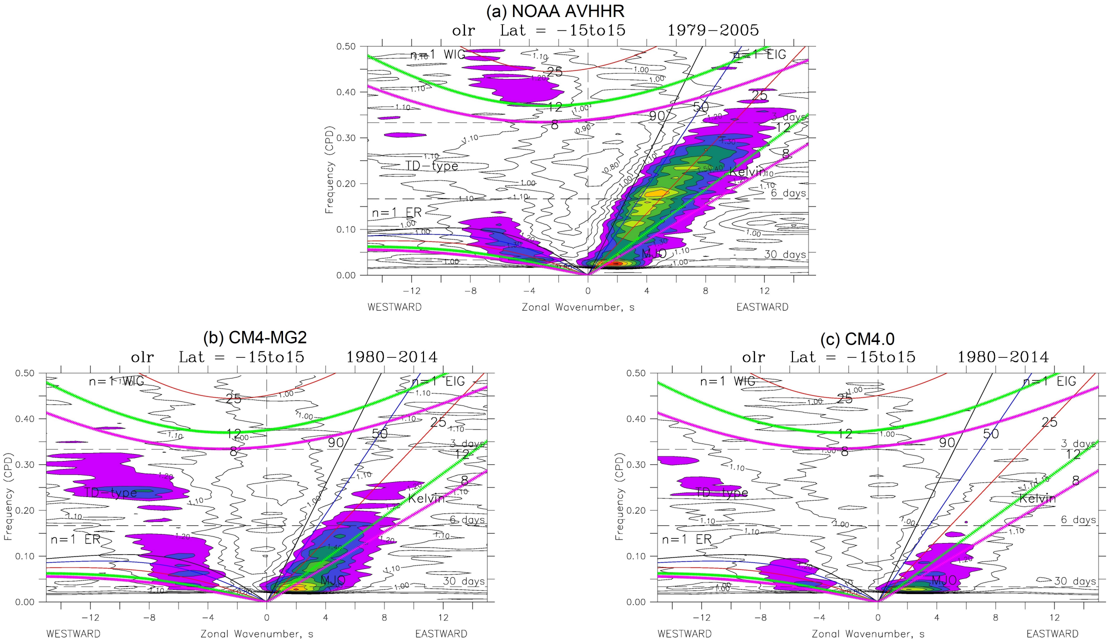

Figure 12. Normalized tropical (15°S–15°N) symmetric power spectra of daily outgoing longwave radiation (OLR): zonal wavenumber versus frequency, from (a) NOAA AVHRR (Advanced Very High Resolution Radiometer) observation, (b) CM4-MG2, and (c) CM4.0. Note that color shading regions of greater than or equal to 1.2 indicate that spectrum power associated MJO, Kelvin, and other convective waves are significant (above background noise).

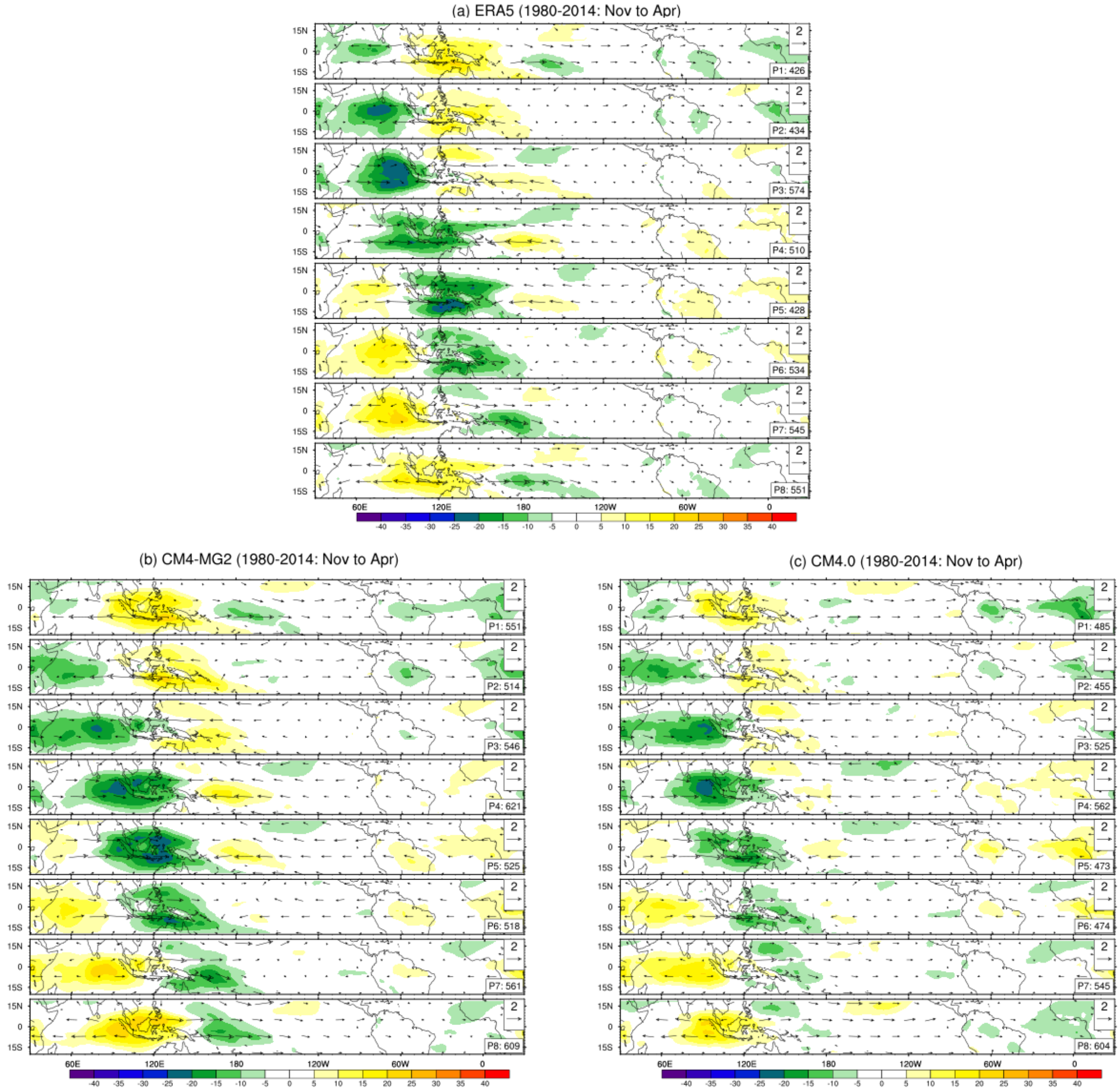

Figure 13. Composites of daily anomalies in OLR (color shaded) and wind vector at 850 hPa (u850, v850) using 20–100 day band-pass filtered data during boreal winter season (November–April) for ERA5 in (a), CM4-MG2 in (b), and CM4.0 in (c).

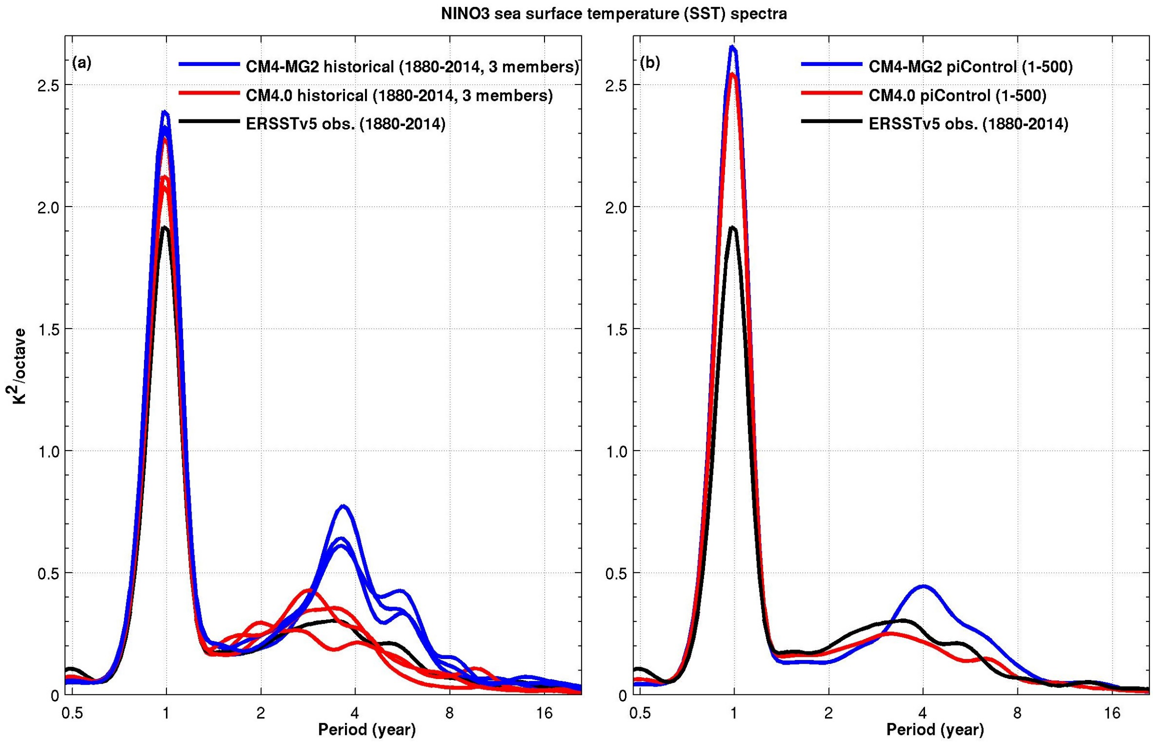

Figure 14. Wavelet power spectra of SST averaged over the Niño-3 region (150–90°W, 5°S–5°N), following Figure 2 of Wittenberg (2009). Black curve is the 1880–2014 time-mean spectrum of the Extended Reconstructed Sea Surface Temperature, version 5 (ERSSTv5) reanalysis (Huang et al., 2017); Colored curves in (a) are the 1880–2014 time-mean spectra for the three CM4-MG2 historical ensemble members (blue), and for the three CM4.0 ensemble members (red). Colored curves in (b) are the time-mean spectra for the corresponding 500-year piControl simulations.

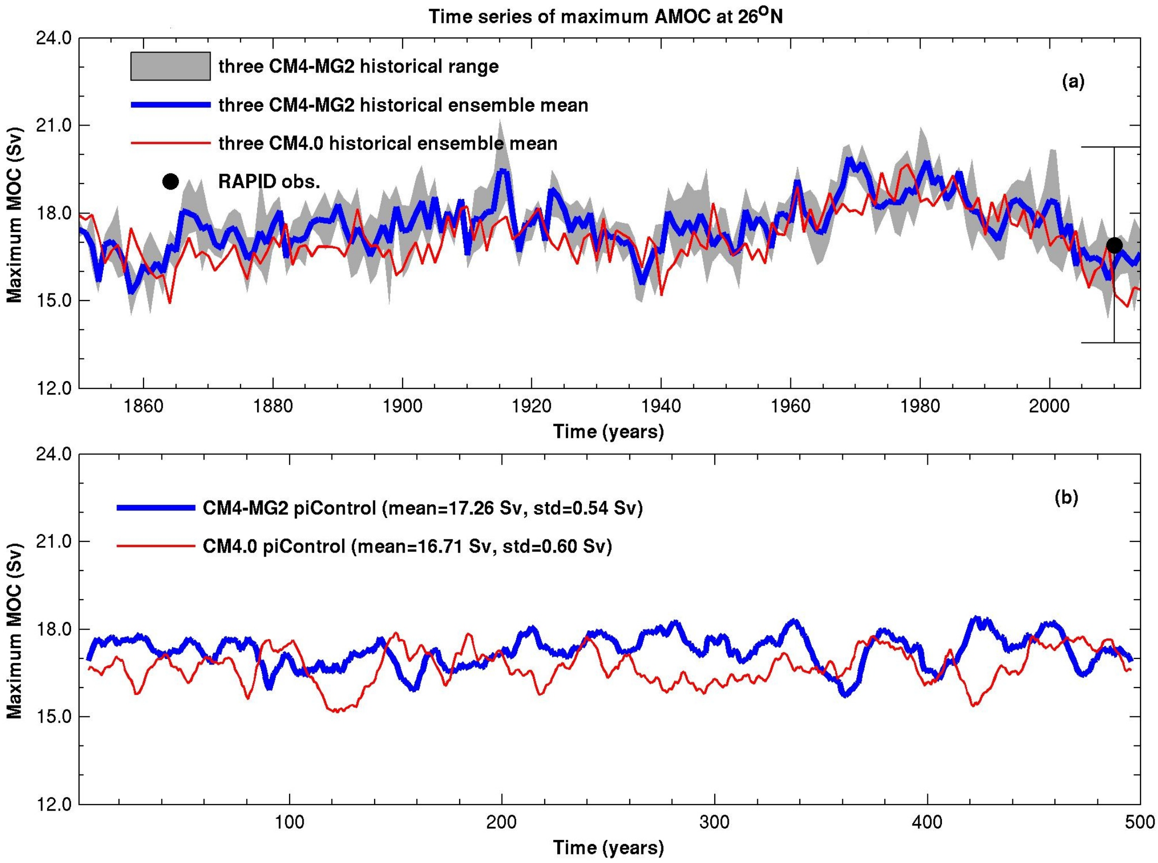

Figure 15. Time evolution of maximum Atlantic Meridional Overturning Circulation (AMOC) at 26°N from three CM4-MG2 historical ensemble mean (thick blue), the spread based on the minimum and maximum values of three CM4-MG2 ensemble members (gray shaded), three CM4-MG2 historical ensemble mean (thin red), and the RAPID array measurement over the period 2004–2015 in (a), and from CM4-MG2 and CM4.0 piControl experiments in (b).

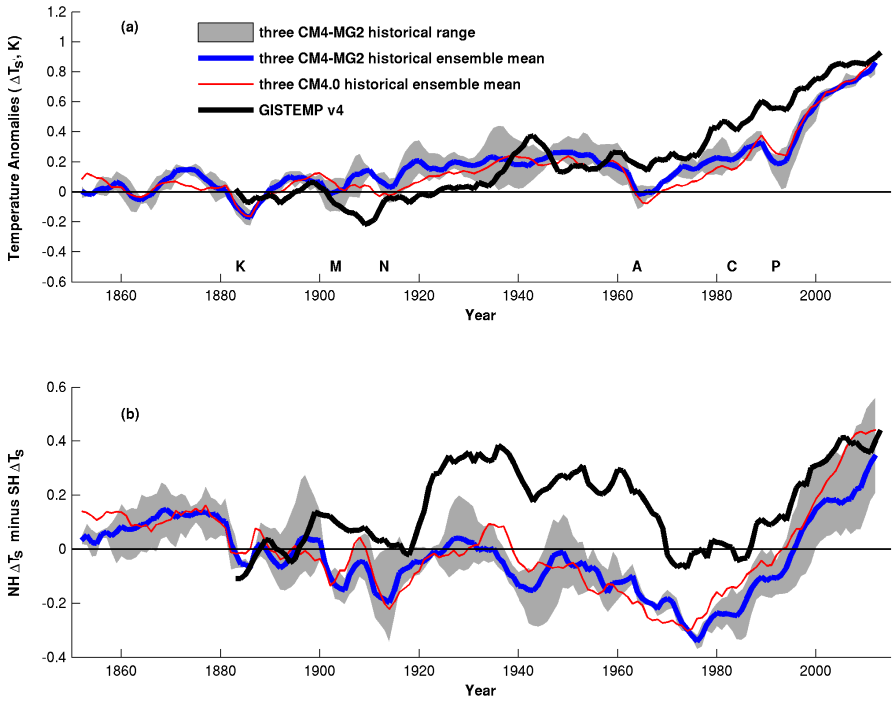

Figure 16. Time series of surface air temperature over land/sea ice and sea surface temperature over open ocean anomalies (ΔT_s, K) from 1880-1900 (a) for the globe, and (b) for the inter-hemispheric contrast between the Northern Hemisphere (NH) and Southern Hemisphere (SH). A 5-year running average is applied to the model results and observations. The observations are from the NASA Goddard Institute for Space Studies Surface Temperature product version 4 (GISTEMP v4) (GISTEMP-Team, 2019; Lenssen et al., 2019). Letters above the horizontal axis mark major volcanic eruptions: Krakatoa (K) in 1883, Santa María (M) in 1902, Novarupta (N) in 1912, Agung (A) in 1963, El Chichón (C) in 1982, and Pinatubo (P) in 1991.

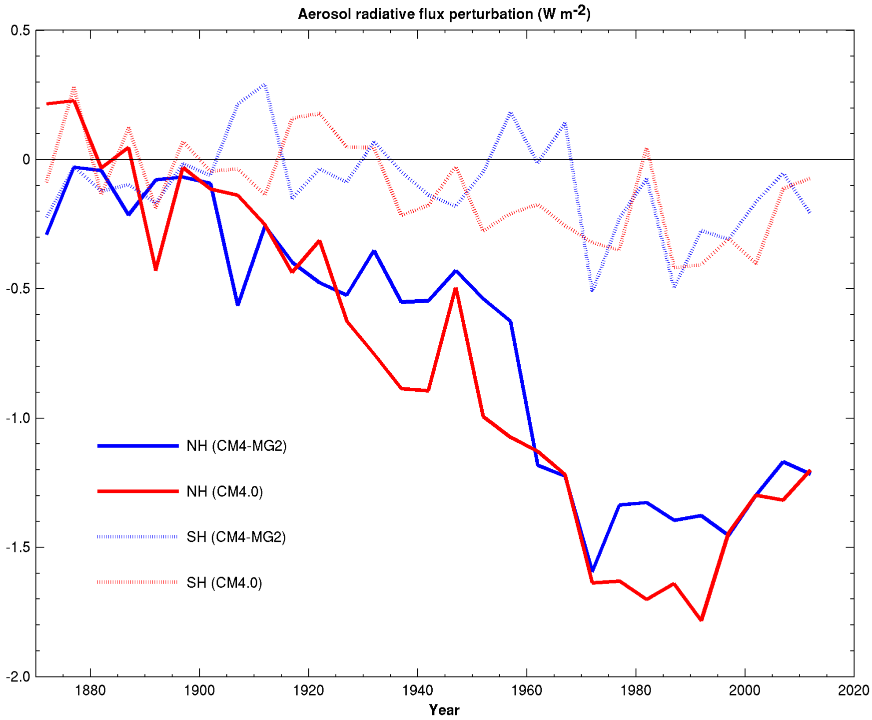

Figure 17. Time series of aerosol radiative flux perturbation (RFP, W m^{-2}) for CM4-MG2 (blue) and CM4.0 (red) in the Northern Hemisphere (NH, solid)) and Southern Hemisphere (SH, dotted) derived from a pair of long AMIP simulations (1870–2014) with prescribed time-varying SST and sea ice concentration. One simulation used the fixed aerosol emission levels at 1850 and the other used the same forcing levels except for time-varying aerosol emissions. Time series are computed by averaging over 5-year period.

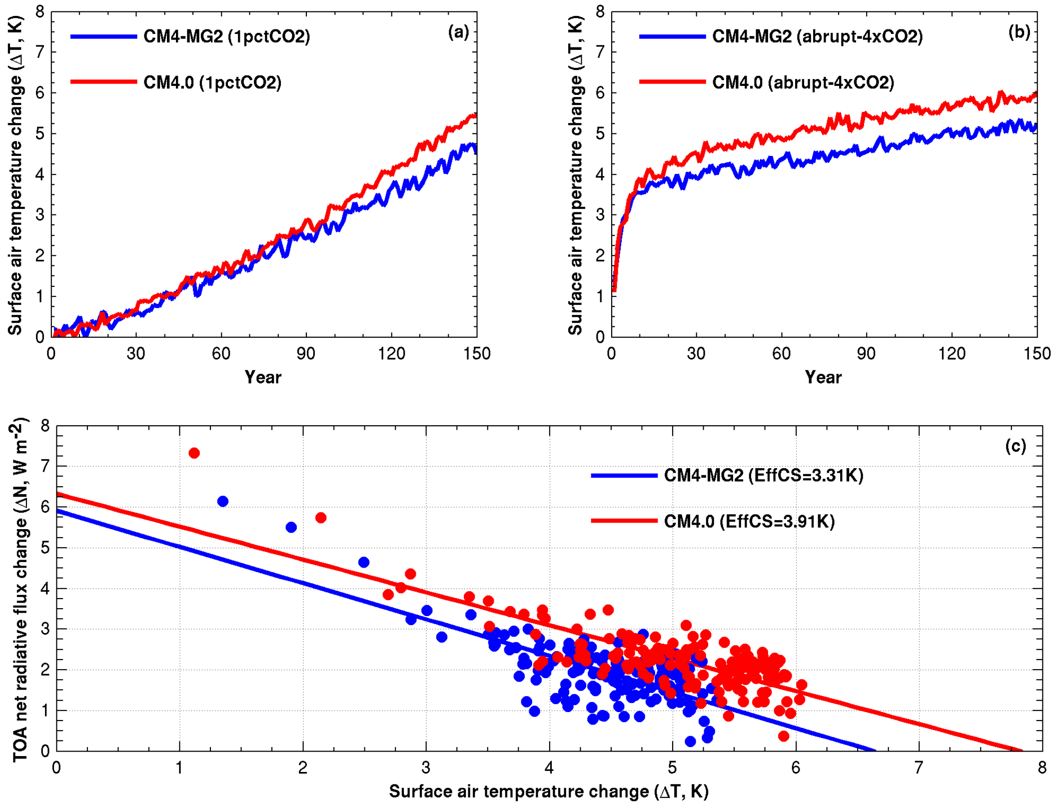

Figure 18. Time series of global annual mean surface air temperature change (ΔT, K) in the 1pctCO2 (a) and abrupt-4xCO2 (b) experiments relative to the pre-industrial control (piControl) experiment. (c) ΔT versus top-of-atmosphere (TOA) net radiative flux change (ΔN, W m^{-2}) of the abrupt-4xCO2 relative to the piControl. Linear regressions are depicted with solid lines for CM4-MG2 (blue) and CM4.0 (red), respectively. The effective climate sensitivity (EffCS) is calculated as the half of the ΔT-axis intercept.

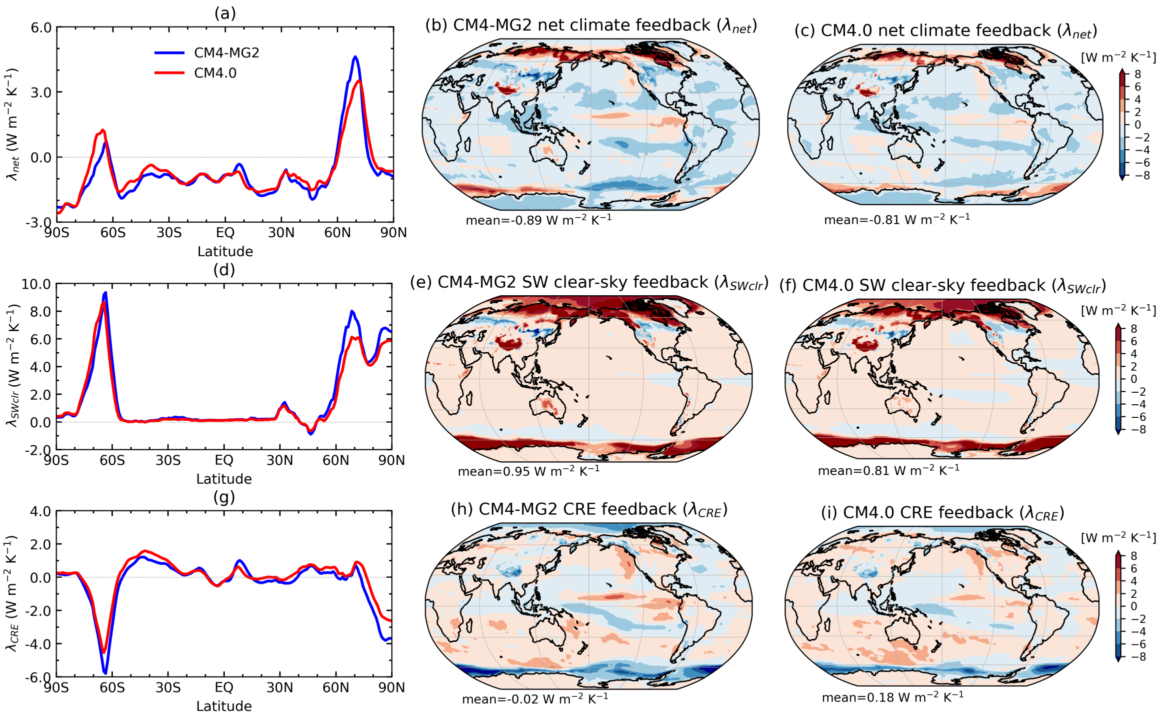

Figure 19. Zonal mean net climate feedback parameter ($\lambda_{\rm net}$, W m^{-2} K^{-1}) in (a), and its shortwave clear-sky component ($\lambda_{\rm SWclr}$) in (d) and cloud radiative effect (CRE) component ($\lambda_{\rm CRE}$) in (g), and their geographical distributions from CM4-MG2 in (b), (e), and (h), and from CM4.0 in (c), (f), and (i). $\lambda_{\rm net}$, $\lambda_{\rm SWclr}$, and $\lambda_{\rm CRE}$ are calculated by regressing the change in net radiative flux at TOA, and its shortwave clear-sky and cloud radiative effect components against surface air temperature change (ΔT) for all 150 years of the abrupt-4xCO2 simulations.

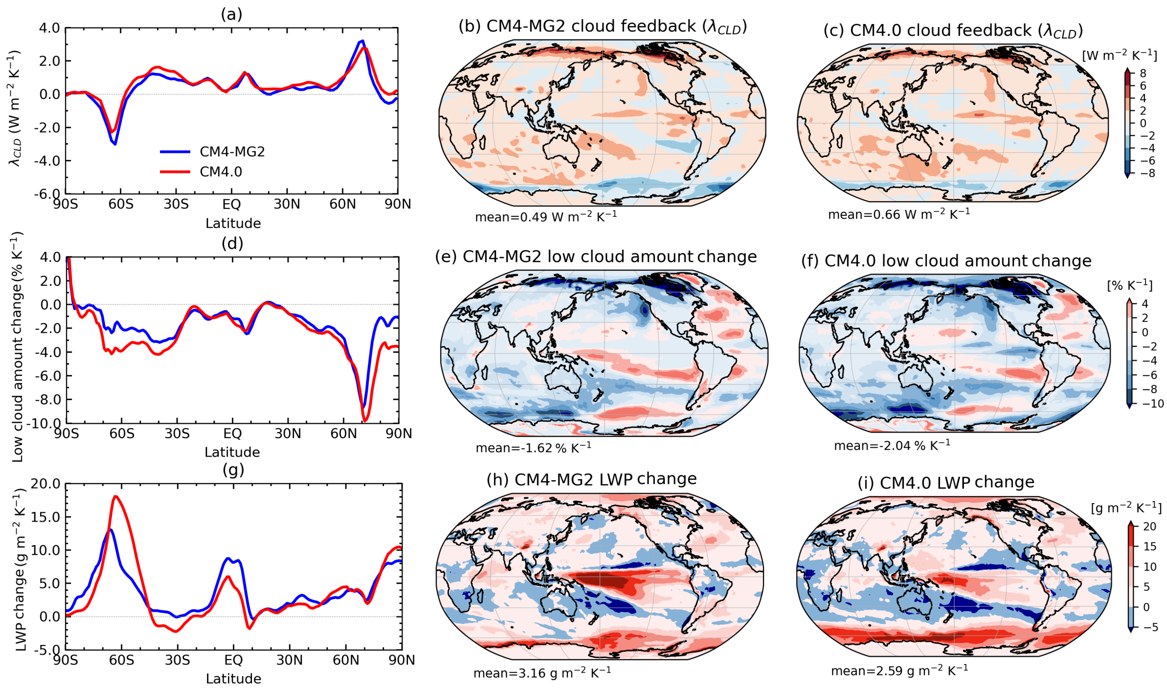

Figure 20. Zonal mean cloud feedback parameter (λ_{CLD}) in (a), low cloud amount feedback in (d), and liquid water path (LWP) feedback in (g), and their geographical distributions from CM4-MG2 in (b), (e), and (h), and from CM4.0 in (c), (f), and (i). λ_{CLD} is estimated using the radiative kernels based on the GFDL model (Soden et al., 2008). The low cloud amount feedback is calculated by regressing the percentage change in low cloud amount against surface air temperature change (ΔT). The liquid water path feedback is calculated by regressing the change in LWP against ΔT.

Table 1. Summary of CM4-MG2 fully coupled atmosphere-ocean-land-sea ice simulations

Experiment	Description	Period (years)	Ensemble size	Initialization
piControl	Pre-industrial control	500	1	piControl spinup
1pctCO2	CO_2 prescribed to increase at 1%/yr	150	1	piControl (101)
abrupt-4xCO2	CO_2 abruptly quadrupled and then held constant	150	1	piControl (101)
historical	Coupled historical	1850-2014	3	piControl (101, 140, 182)

.

Table 2. Global-annual means of three ensemble members of CM4-MG2 and CM4.0 historical simulations for 1980-2014, and observations: shortwave cloud radiative effect (SWCRE), longwave cloud radiative effect (LWCRE), clear-sky shortwave absorption (SWABS_clr), clear-sky outgoing longwave radiation (OLR_clr) at TOA based on the CERES-EBAF (Loeb et al., 2018); surface latent heat flux (LH_flx), surface sensible heat fluxes (SH_flx), surface zonal wind stress (tau_x), surface meridional wind stress (tau_y), sea level pressure in the Northern Hemisphere (SLP_NH), sea level pressure in the Southern Hemisphere (SLP_SH) based on the ERA-Interim reanalyses (Dee et al., 2011); convective and stratiform liquid water path (LWP_{cw}) over ocean based on the Multi-Sensor Advanced Climatology of Liquid Water Path (MAC-LWP) (Elsaesser et al., 2017); and total ice water path (IWP_{tot}) based on the CloudSat (J. Jiang et al., 2012). Values in parenthesis indicate root-mean-square-errors (RMSEs).

Variable	Observations	CM4-MG2	CM4.0
SWCRE (W m^{-2})	45.39	-48.71 (9.28)	-48.74 (9.48)
LWCRE (W m^{-2})	25.89	22.78 (5.32)	23.65 (5.18)
SWABS_clr (W m^{-2})	286.93	287.37 (7.10)	287.38 (7.20)
OLR_clr (W m^{-2})	268.22	260.80 (8.35)	261.48 (7.70)
LH_flx (W m^{-2})	83.17	82.40 (9.45)	83.64 (9.48)
SH_flx (W m^{-2})	17.48	18.49 (6.34)	18.22 (6.33)
tau_x (dPa)	0.12	0.08 (0.18)	0.09 (0.20)
tau_y (dPa)	0.02	0.004 (0.17)	0.002 (0.17)
SLP_NH (hPa)	1013.62	1013.16 (1.12)	1013.25 (1.00)
SLP_SH (hPa)	1009.02	1007.95 (2.57)	1007.72 (2.70)
[1]LWP_{cw} ocean (g m^{-2})	81.06	80.53 (16.85)	60.50 (28.78)
[2]IWP_{tot} (g m^{-2})	70.14	53.90 (39.78)	52.55 (40.73)

[1] LWP_{cw} includes both stratiform and convective cloud water, but not rain.
[2] IWP_{tot} includes stratiform and convective cloud ice and snow.

Table 3. Global mean CO_2 effective radiative forcing, sensitivity, and feedback due to CO_2 doubling

	CM4-MG2	CM4.0
TCR (K)	1.85	2.05
EffCS (K)	3.31	3.91
EffF_{2x} (W m^{-2})	2.95	3.16
λ_{net} (W m^{-2} K^{-1})	-0.89	-0.81
λ_{SWclr} (W m^{-2} K^{-1})	0.95	0.81
λ_{LWclr} (W m^{-2} K^{-1})	-1.82	-1.80
λ_{CRE} (W m^{-2} K^{-1})	-0.02	0.18
λ_{SWCRE} (W m^{-2} K^{-1})	-0.20	-0.06
λ_{LWCRE} (W m^{-2} K^{-1})	0.18	0.24
λ_{CLD} (W m^{-2} K^{-1})	0.49	0.66
λ_{SWCLD} (W m^{-2} K^{-1})	-0.04	0.09
λ_{LWCLD} (W m^{-2} K^{-1})	0.53	0.57
$\lambda_{\mathrm{albedo}}$ (W m^{-2} K^{-1})	0.48	0.47
$\lambda_{\mathrm{Planck}}$ (W m^{-2} K^{-1})	-3.53	-3.55
λ_{LR} (W m^{-2} K^{-1})	-0.25	-0.20
λ_{vapor} (W m^{-2} K^{-1})	1.73	1.68
Cess feedback (W m^{-2} K^{-1})	-2.02	-1.77

TCR (transient climate response) is global mean surface air temperature change (ΔT, K) at the time of doubled CO_2 (Year 70) in the 1pctCO2 experiment (Table 1), evaluated as a time-mean over years 61–80 (J. M. Gregory & Forster, 2008). EffCS, EffF_{2x}, and λ_{net} are the effective climate sensitivity, 2xCO_2 radiative forcing, and net climate feedback parameter, respectively. They are estimated from a linear regression of net radiative flux change (ΔN, W m^{-2}) at top-of-atmosphere (TOA) against ΔT for all 150 years of the abrupt-4xCO_2 experiment (Table 1). EffCS and EffF_{2x} are the ΔT-axis and ΔN-axis intercepts divided by 2; λ_{net} is the slope of the linear regression line. λ_{SWclr}, λ_{LWclr}, and λ_{CRE} are clear-sky shortwave, clear-sky longwave, and cloud radiative effect (CRE) feedback parameters. The total feedback is also decomposed into cloud (λ_{CLD}), surface albedo ($\lambda_{\mathrm{albedo}}$), Planck ($\lambda_{\mathrm{Planck}}$), lapse rate ($\lambda_{\mathrm{LR}}$), and water vapor ($\lambda_{\mathrm{vapor}}$) feedback components using the radiative kernels based on the GFDL AM2 model (Soden et al., 2008). The cloud feedback (λ_{CLD}) is further decomposed into shortwave (λ_{SWCLD}) and longwave (λ_{LWCLD}) cloud feedback. The Cess feedback is calculated as ΔN divided by the warming of sea surface temperature (SST) from a pair of present-day simulation and global warming simulation with SST uniformly increased by 2 K.

Data Availability Statement

The CM4.0 source codes are available at https://doi.org/10.5281/zenodo.3339397. The CM4.0 model data have been deposited in the CMIP6 archive with the identifier https://doi.org/10.22033/ESGF/CMIP6.1402 and https://doi.org/10.22033/ESGF/CMIP6.8594.

The original MG2 source code was from the CESM2.1.3 release, which can be downloaded at http://www.cesm.ucar.edu/models/cesm2/release_download.html. The CM4-MG2 source codes can be found at https://doi.org/10.5281/zenodo.6323646. The CM4-MG2 model data is available at ftp://data1.gfdl.noaa.gov/users/huan.guo/microphysics/CM4-MG2.

The radiative kernels for calculating the cloud feedback are accessible via https://climate.rsmas.miami.edu/data/radiative-kernels/index.html. The CERES-EBAF and GPCP data can be obtained from https://ceres.larc.nasa.gov/data and https://psl.noaa.gov/data/gridded/data.gpcp.html, respectively. The GISS Surface Temperature Analysis (GISTEMP v4) is accessible via https://data.giss.nasa.gov/gistemp. The HadISST data set can be downloaded at https://www.metoffice.gov.uk/hadobs/hadisst/data/download.html. The NOAA Extended Reconstructed Sea Surface Temperature (SST) V5 is available at https://psl.noaa.gov/data/gridded/data.noaa.ersst.v5.html. Data from the RAPID AMOC monitoring project are freely available from www.rapid.ac.uk/rapidmoc (https://doi.org/10.5285/aa57e879-4cca-28b6-e053-6c86abc02de5). The ERA-Interim (European Center for Medium Range Weather Forecasting Re-Analysis Interim) and ERA5 data are available at https://www.ecmwf.int/en/research/climate-reanalysis/era-interim and https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5, respectively.

Acknowledgments

We thank Shian-Jiann Lin for designing and sharing the time-implicit sedimentation scheme. Mitchell Bushuk and David Paynter are acknowledged for their comments and suggestions on an internal review of the manuscript.

References

Adames, Á. F., & Wallace, J. M. (2015). Three-dimensional structure and evolution of the moisture field in the MJO. *Journal of the Atmospheric Sciences*(72), 3733–3754. doi: 10.1175/JAS-D-15-0003.1

Adcroft, A., Anderson, W., Balaji, V., Blanton, C., Bishuk, M., et al. (2019). The GFDL global ocean and sea ice model OM4.0: Model description and simulation features. *Journal of Advances in Modeling Earth Systems*. doi: 10.1029/2019MS001726

Adcroft, A., & Hallberg, R. (2006). On methods for solving the oceanic equations of motion in generalized vertical coordinates. *Ocean Modelling*, *11*(1), 224—233. doi: 10.1016/j.ocemod.2004.12.007

Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P., Janowiak, J., ... Arkin, P. (2003). The Version 2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979-Present). *J. Hydrometeor.*, *4*, 1147–1167.

Adler, R. F., Sapiano, M., Huffman, G., Bolvin, D., Gu, G., Wang, J., ... Becker, A. (2016). *The new version 2.3 of the Global Precipitation Climatology Project (GPCP) monthly analysis product* (Internal Report). College Park, MD: Earth System Science Interdisciplinary Center, Univ. of Maryland.

Alexander, M. J., & Dunkerton, T. J. (1999). A spectral parameterization of mean-flow forcing due to breaking gravity waves. *Journal of the Atmospheric Sciences*, *56*(24), 4167-–4182. doi: 10.1175/1520-0469(1999)056⟨4167:ASPOMF⟩2.0.CO;2

Andrews, T., Andrews, M. B., Bodas-Salcedo, A., Jones, G. S., Kuhlbrodt, T., Manners, J., & et al. (2019). Forcing, feedbacks and climate sensitivity in HadGEM3-GC3.1 and UKESM1. *Journal of Advances in Modeling Earth Systems*, *11*, 4377–4394. doi: 10.1029/2019MS001866

Andrews, T., Gregory, J. M., Webb, M. J., & Taylor, K. E. (2012). Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models. *Geophysical Research Letters*, *39*(L09712). doi: 10.1029/2012GL051607

Antonov, J. I., Locarnini, R. A., Boyer, T. P., Mishonov, A. V., & Garcia, H. E. (2006). World Ocean Atlas 2005, Volume 2: Salinity. . In *S. levitus (ed.), NOAA Atlas NESDIS 62* (pp. 1–182). Washington, D.C.: U.S. Government Printing Office.

Benedict, J. J., Maloney, E. D., Sobel, A. H., & Frierson, D. M. W. (2014). Gross moist stability and MJO simulation skill in three full-physics GCMs. *Journal of the Atmospheric Sciences*(71), 3327—3349. doi: 10.1175/JAS-D-13-0240.1

Bleck, R. (2002). An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates . *Ocean Modelling*, *4*(1), 55–88. doi: 10.1016/S1463-5003(01)00012-9

Bodas-Salcedo, A., Mulcahy, J. P., Andrews, T., Williams, K. D., Ringer, M. A., Field, P. R., & Elsaesser, G. S. (2019). Strong dependence of atmospheric feedbacks on mixed-phase microphysics and aerosol-cloud interactions in HadGEM3. *Journal of Advances in Modeling Earth Systems*, *11*(6), 1735–1758. doi: 10.1029/2019MS001688

Bogenschutz, P. A., Yamaguchi, T., & Lee, H.-H. (2021). The Energy Exascale Earth System Model simulations with high vertical resolution in the lower troposphere. *Journal of Advances in Modeling Earth Systems*, *13*(e2020MS002239). doi: 10.1029/2020MS002239

Brient, F., Schneider, T., Tan, Z., Bony, S., Qu, X., & Hall, A. (2015). Shallowness of tropical low clouds as a predictor of climate models' response to warming. *Climate Dyn.*, *47*, 433–449. doi: 10.1007/s00382-015-2846-0

Cavalieri, D. J., Parkinson, C. L., Gloersen, P., & Zwally, H. J. (1996). Sea ice concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data, Version 1. *NASA DAAC at the Natl. Snow and Ice Data Cent.*. doi: 10.5067/8GQ8LZQVL0VL

Cess, R. D., & Coauthors. (1990). Intercomparison and interpretation of climate feedback processes in 19 Atmospheric General Circulation Models. *Journal of Geophysical Research*, *95*(D10), 16,601-16,615.

Danabasoglu, G., Lamarque, J., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., … et al (2020). The Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*(2). doi: 10.1029/2019MS001916

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., … Vitart, F. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, *137*, 553–597. doi: doi:10.1002/qj.828

Delworth, T. L., Broccoli, A. J., Rosati, A., Stouffer, R. J., Balaji, V., Beesley, J. A., … et. al. (2006). GFDL's CM2 global coupled climate models. Part I: Formulation and simulation characteristics. *Journal of Climate*, *19*(5), 643–674. doi: 10.1175/JCLI3629.1

Delworth, T. L., Cooke, W. F., Adcroft, A., Bushuk, M., Chen, J. H., Dunne, K. A., & et al. (2020). Spear: The next generation GFDL modeling system for seasonal to multidecadal prediction and projection. *Journal of Advances in Modeling Earth Systems*, *12*(e2019MS001895). doi: 10.1029/2019MS001895

Delworth, T. L., & Dixon, K. W. (2006). Have anthropogenic aerosols delayed a greenhouse gas-induced weakening of the North Atlantic thermohaline circulation? *Geophysical Research Letters*, *33*(2).

Donner, L. J., Wyman, B. L., Hemler, R. S., Horowitz, L. W., Ming, Y., Zhao, M., … others (2011). The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL global coupled model CM3. *Journal of Climate*, *24*(13), 3484–3519.

Dunne, J. P., Horowitz, L. W., Adcroft, A. J., Ginoux, P., Held, I. M., John, J. G., & et al. (2020). The GFDL Earth System Model Version 4.1 (gfdl-esm 4.1): Overall coupled model description and simulation characteristics. *Journal of Advances in Modeling Earth Systems*, *12*(e2019MS002015). doi: 10.1029/2019MS002015

Dunne, J. P., Winton, M., Bacmeister, J., Danabasoglu, G., Gettelman, A., Golaz, J., & et al. (2020). Comparison of equilibrium climate sensitivity estimates from slab ocean, 150-year, and longer simulations. *Geophysical Research Letters*, *47*(e2020GL088852). doi: 10.1029/2020GL088852

Elsaesser, G. S., O'Dell, C. W., Lebsock, D., M., Bennartz, R., Greenwald, T. J., & Wentz, F. J. (2017). The Multi-Sensor Advanced Climatology of Liquid Water Path (MAC-LWP). *J. Climate*, *30*(24), 10193–10210. doi: doi:10.1175/JCLI-D-16-0902.1

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

Fan, S., Ginoux, P., Seman, C. J., Silvers, L. G., & Zhao, M. (2019). Toward improved cloud-phase simulation with a mineral dust and temperature-dependent parameterization for ice nucleation in mixed-phase clouds. *Journal of the Atmospheric Sciences*, *76*(11). doi: 10.1175/JAS-D-18-0287.1

Fan, S., Knopf, D. A., Heymsfield, A. J., & Donner, L. J. (2017). Modeling of aircraft measurements of ice crystal concentration in the Arctic and a parameterization for mixed-phase cloud. *Journal of the Atmospheric Sciences*, *74*, 3799–3814. doi: 10.1175/JAS-D-17-0037.1

Fels, S., & Schwarzkopf, D. M. (1975). The simplified exchange approximation: A new method for radiative transfer calculations. *Journal of the Atmospheric*

Sciences, *32*, 1475–1488.

Forster, P. M., Richardson, T., Maycock, A. C., Smith, C. J., Samset, B. H., Myhre, G., ... Schulz, M. (2016). Recommendations for diagnosing effective radiative forcing from climate models for CMIP6. *Journal of Geophysical Research: Atmospheres*, *121*, 12,460–12,475. doi: 10.1002/2016JD025320

Freidenreich, S. M., & Ramaswamy, V. (2005). Refinement of the Geophysical Fluid Dynamics Laboratory solar benchmark computations and an improved parameterization for climate models. *Journal of Geophysical Research*, *110*(D17105). doi: 10.1029/2004JD005471

Fu, Q., Krueger, S. K., & Liou, K. N. (1995). Interactions between radiation and convection in simulated tropical cloud clusters . *Journal of the Atmospheric Sciences*, *52*(9), 1310–1328.

Garner, S. T. (2005). A topographic drag closure built on an analytical base flux. *Journal of the atmospheric sciences*, *62*(7), 2302–2315.

Garner, S. T. (2018). Ground-truth model evaluation of subgrid orographic base-flux parameterization. *Journal of the Atmospheric Sciences*, *75*(10), 3653–3670.

Gettelman, A., & Morrison, H. (2015a). Advanced two-moment bulk microphysics for global models. Part I: Off-line tests and comparison with other schemes. *J. Climate*, *28*(3), 1268–1287. doi: 10.1175/JCLI-D-14-00103.1

Gettelman, A., Morrison, H., Santos, S., Bogenshutz, P., & Caldwell, P. M. (2015b). Advanced two-moment bulk microphysics for global models. Part II: Global model solutions and aerosol-cloud interactions. *J. Climate*, *28*(3), 1288–1306. doi: 10.1175/JCLI-D-14-00102.1

Gettelman, A., Morrison, H., Terai, C. R., & Wood, R. (2013). Microphysical process rates and global aerosol–cloud interactions. *Atmos. Chem. Phys.*, *13*(11), 11 789–11 825. doi: 10.5194/acpd-13-11789-2013

GFDL Global Atmosphere Model Development Team. (2004). The new GFDL global atmosphere and land model AM2-LM2: Evaluation with prescribed SST simulations. *J. Climate*, *17*, 4641–4673. doi: 10.1175/JCLI-3223.1

GISTEMP-Team. (2019). GISS Surface Temperature Analysis (GISTEMP), version 4. NASA Goddard Institute for Space Studies. , *Dataset accessed 2016-09-27 at https://data.giss.nasa.gov/gistemp/*.

Golaz, J.-C., Caldwell, P. M., Van Roekel, L. P., Petersen, M. R., Tang, Q., Wolfe, J. D., & et al. (2019). The DOE E3SM coupled model version 1:overview and evaluation at standard resolution. *Journal of Advances in Modeling Earth Systems*, *11*, 2089–2129. doi: 10.1029/2018MS001603

Golaz, J.-C., Horowitz, L. W., & Levy, H. (2013). Cloud tuning in a coupled climate model: Impact on 20th century warming. *Geophysical Research Letters*, *40*(10), 2246–2251.

Golaz, J.-C., Salzmann, M., Donner, L. J., Horowitz, L. W., Ming, Y., & Zhao, M. (2011). Sensitivity of the aerosol indirect effect to subgrid variability in the cloud parameterization of the gfdl atmosphere general circulation model AM3. *J. Climate*, *24*, 3145–3160. doi: 10.1175/2010JCLI3945.1

Gregory, J., Ingram, W., Palmer, M., Jones, G., Stott, P., Thorpe, R., ... Williams, K. (2004). A new method for diagnosing radiative forcing and climate sensitivity. *Geophysical Research Letters*, *31*(3).

Gregory, J. M., & Forster, P. M. (2008). Transient climate response estimated from radiative forcing and observed temperature change. *Journal of Geophysical Research*, *113*(D23105). doi: 10.1029/2008JD010405

Guilyardi, E., Capotondi, A., Lengaigne, M., Thual, S., & Wittenberg, A. T. (2020). ENSO modeling: History, progress, and challenges. In McPhaden, M J and A Santoso and W Cai (Ed.), (pp. 201–226). Washington, DC: American Geophysical Union. doi: 10.1002/9781119548164.ch9

Guo, H., Ming, Y., Fan, S., Zhou, L., Harris, L., & Zhao, M. (2021). Two-moment bulk cloud microphysics with prognostic precipitation in GFDL's atmosphere model AM4.0: Configuration and performance. *Journal of Advances in Modeling Earth Systems*, *13*(e2020MS002453). doi: 10.1029/2020MS002453

Hansen, J., Sato, M., Ruedy, R., Nazarenko, L., Lacis, A., Schmidt, G. A., … et al. (2014). Efficacy of climate forcings. *Journal of Geophysical Research*, *110*(D18104). doi: 10.1029/2005jd005776

Harris, I., Jones, P., Osborn, T., & Lister, D. (2014). Updated high-resolution grids of monthly climatic observations—The CRU TS3.10 dataset. *International Journal of Climatology*, *34*, 623–642.

Harris, L., Zhou, L., Chen, X., & Chen, J.-H. (2020). *The GFDL Finite-Volume Cubed-Sphere Dynamical Core* (NOAA Technical Memorandum). Princeton, NJ: OAR GFDL. doi: 10.25923/7h88-c534

Hassan, T., Allen, R. J., Liu, W., & Randles, C. A. (2021). Anthropogenic aerosol forcing of the atlantic meridional overturning circulation and the associated mechanisms in CMIP6 models. *Atmos. Chem. Phys.*, *21*, 5821–5846. doi: 10.5194/acp-21-5821-2021

Held, I. M., Guo, H., Adcroft, A., Dunne, J. P., Horowitz, L. W., Krasting, J., & et al. (2019). Structure and performance of GFDL's CM4.0 climate model. *Journal of Advances in Modeling Earth Systems*, *11*(11), 3691–3727. doi: 10.1029/2019MS001829

Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., … Zhang, H. (2017). Extended Reconstructed Sea Surface Temperature, version 5 (ERSSTv5): Upgrades, validations, and intercomparisons. *J. Climate*, *30*, 8179–8205. doi: 10.1175/JCLI-D-16-0836.1

Jakob, C., & Klein, S. A. G. (2000). A parameterization of the effects of cloud and precipitation overlap for use in general circulation models. *Q. J. Roy. Meteorol. Soc.*, *126*, 2525–2544.

Jiang, J., Su, H., Zhai, C., Perun, V., & et al. (2012). Evaluation of Cloud and Water Vapor Simulations in CMIP5 Climate Models Using NASA A-Train Satellite Observations. *J. Geophys. Res.*, *117*. doi: 10.1029/2011JD017237

Jiang, X. (2017). Key processes for the eastward propagation of the Madden-Julian Oscillation based on multimodel simulations. *J. Geophys. Res. Atmos.*, *122*(755–770). doi: 10.1002/2016JD025955

Jiang, X., Adames, Á. F., Kim, D., Maloney, E. D., Lin, H., Kim, H., & et al. (2020). Fifty years of research on the Madden-Julian Oscillation: Recent progress, challenges, and perspectives . *J. Geophys. Res. Atmos.*, *125*(e2019JD030911). doi: 10.1029/2019JD030911

Kelley, M., Schmidt, G. A., Nazarenko, L., Bauer, S. E., Ruedy, R., Russell, G. L., … et al. (2020). GISS-E2.1: Configurations and climatology. *J. Adv. Model. Earth Syst.*, *12*(8). doi: 10.1029/2019MS002025

Kessler, W. S. (2002). Is ENSO a cycle or a series of events? *Geophysical Research Letters*, *29*(23). doi: 10.1029/2002GL015924

Kim, H.-M. (2017). The impact of the mean moisture bias on the key physics of MJO propagation in the ECMWF reforecast. *J. Geophys. Res. Atmos.*, *122*, 7772–7784. doi: 10.1002/2017JD027005

Kim, W. M., Yeager, S., Chang, P., & Danabasoglu, G. (2017). Low-frequency North Atlantic climate variability in the Community Earth System Model large ensemble. *Journal of Climate*, *31*, 787-–813. doi: 10.1175/JCLI-D-17-0193.1

Klein, S. A., & Hartmann, D. L. (1993). The seasonal cycle of low stratiform clouds. *J. Climate*, *6*, 1587-1606.

Large, W. G., & Danabasoglu, G. (2006). Attribution and impacts of upper-ocean biases in CCSM3. *J. Climate*, *19*(11), 2325–2345. doi: 10.1175/JCLI3740.1

Larkin, N. K., & Harrison, D. E. (2002). ENSO warm (El Niño) and cold (La Niña) event life cycles: Ocean surface anomaly patterns, their symmetries, asymmetries, and implications. *J. Climate*, *15*, 1118-–1140. doi: 10.1175/1520-0442(2002)015⟨1118:EWENOA⟩2.0.CO;2

Lee, H.-H., Bogenschutz, P., & Yamaguchi, T. (2021). The implementation of framework for improvement by vertical enhancement into Energy Exascale Earth System Model. *Journal of Advances in Modeling Earth Systems*, *13*(e2020MS002240). doi: 10.1029/2020MS002240

Lenssen, N., Schmidt, G., Hansen, J., Menne, M., Persin, A., Ruedy, R., & Zyss, D. (2019). Improvements in the GISTEMP Uncertainty Model. *Journal of Geophysical Research: Atmospheres*, *124*. doi: https://doi.org/10.1029/2018JD029522

Liebmann, B., & Smith, C. A. (1996). Description of a complete (interpolated) Outgoing Longwave Radiation dataset. *Bulletin of the American Meteorological Society*, *77*, 1275–1277.

Lin, S.-J. (2004). A "vertically Lagrangian" finite-volume dynamical core for global models. *Monthly Weather Review*, *132*(10), 2293–2307. doi: 10.1175/1520-0493(2004)132⟨2293:AVLFDC⟩2.0.CO;2

Locarnini, R. A., Mishonov, A. V., Antonov, J. I., Boyer, T. P., & Garcia, H. E. (2006). World Ocean Atlas 2005, Volume 1: Temperature. In *S. levitus (ed.), NOAA Atlas NESDIS 61* (pp. 1–182). Washington, D.C.: U.S. Government Printing Office.

Lock, A. P., Brown, A. R., Bush, M. R., Martin, G. M., & Smith, R. N. B. (2000). A new boundary layer mixing scheme. Part I: scheme description and single-column model tests. *Mon. Wea. Rev.*, *128*, 3187–3199.

Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J. G., … Kato, S. (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.

Loeb, N. G., Wielicki, B. A., Doelling, D. R., Smith, G. L., Keyes, D. F., Kato, S., … Wong, T. (2009). Toward optimal closure of the Earth's top-of-atmosphere radiation budget. *Journal of Climate*, *22*(3), 748–766.

Lohmann, U., Rotstayn, L., Storelvmo, T., Jones, A., Menon, S., Quaas, J., … Ruedy, R. (2010). Total aerosol effect: Radiative forcing or radiative flux perturbation? *Atmos. Chem. Phys.*, *10*, 3235–3246. doi: 10.5194/acp-10-3235-2010

Madden, R. A., & Julian, P. R. (1971). Detection of a 40-50 day oscillation in the zonal wind in the tropical pacific. *J. Atmos. Sci.*, *28*, 702–708.

Madden, R. A., & Julian, P. R. (1972). Description of global-scale circulation cells in the tropics with a 40-50 day period. *J. Atmos. Sci.*, *29*, 1109–1123.

McPhaden, M J and A Santoso and W Cai (Ed.). (2020). *El Niño Southern Oscillation in a Changing Climate.* Washington, DC: American Geophysical Union. doi: 10.1002/9781119548164

Menary, M. B., Robson, J., Allan, R. P., Booth, B. B., Cassou, G., C.and Gastineau, Gregory, J., … Zhang, R. (2020). Aerosol-forced AMOC changes in CMIP6 historical simulations. *Geophysical Research Letters*, *47*(14). doi: 10.1029/2020GL088166

Meyers, M. P., DeMott, P. J., & Cotton, W. R. (1992). New primary ice nucleation parameterizations in an explicit cloud model. *J. Appl. Meteorol.*, *32*, 708–721.

Milly, P. C., Malyshev, S. L., Shevliakova, E., Dunne, K. A., Findell, K. L., Gleeson, T., … Swenson, S. (2014). An enhanced model of land water and energy for global hydrologic and earth-system studies. *Journal of Hydrometeorology*, *15*(5), 1739–1761.

Ming, Y., Ramaswamy, V., Donner, L. J., & Phillips, V. T. J. (2006). A new parameterization of cloud droplet activation applicable to General Circulation Models. *J. Atmos. Sci.*, *63*, 1348-1356.

Ming, Y., Ramaswamy, V., Donner, L. J., Phillips, V. T. J., Klein, S. A., Ginoux, P. A., & Horowitz, L. W. (2007). Modeling the interactions between aerosols and liquid water clouds with a self-consistent cloud scheme in a General Circulation Model. *J. Atmos. Sci.*, *64*(4), 1189-1209. doi: 10.1175/JAS3874.1

Moat, B. I., Frajka-Williams, E., Smeed, D., Rayner, D., Sanchez-Franks, A., Johns, W. E., … Collins, J. (2020). Atlantic meridional overturning circulation observed by the RAPID-MOCHA-WBTS (RAPID-Meridional Overturning Circulation and Heatflux Array-Western Boundary Time Series) array at 26N from 2004 to 2018. *(v2018.2)*. doi: 10.5285/57e879-4cca-28b6-e053-6c86abc02de5

Mülmenstädt, J., Nam, C., Salzmann, M., Kretzschmar, J., L'Ecuyer, T. S., Lohmann, U., … Quaas, J. (2020). Reducing the aerosol forcing uncertainty using observational constraints on warm rain processes. *Science Advances*, *6*(22), eaaz6433. doi: 10.1126/sciadv.aaz6433

Paynter, D. J., & Ramaswamy, V. (2012). Variations in water vapor continuum radiative transfer with atmospheric conditions. *Journal of Geophysical Research*, *117*(D16310). doi: 10.1029/2012JD017504

Paynter, D. J., & Ramaswamy, V. (2014). Investigating the impact of the shortwave water vapor continuum upon climate simulations using GFDL global models. *Journal of Geophysical Research: Atmospheres*, *119*, 10720–10737. doi: 10.1002/2014JD021881

Posselt, R., & Lohmann, U. (2008). Introduction of prognostic rain in ECHAM5: Design and single column model simulations. *Atmos. Chem. Phys.*, *8*, 2949–2963. doi: doi:10.5194/acp-8-2949-2008

Posselt, R., & Lohmann, U. (2009). Sensitivity of the total anthropogenic aerosol effect to the treatment of rain in a global climate model. *Geophys. Res. Lett.*, *36*(L02805). doi: doi:10.1029/2008GL035796

Pritchard, M. S., & Bretherton, C. S. (2014). Causal evidence that rotational moisture advection is critical to the superparameterized Madden–Julian Oscillation. *Journal of the Atmospheric Sciences*, *71*, 800–815. doi: doi:10.1175/JAS-D-13-0119.1

Putman, W. M., & Lin, S.-J. (2007). Finite-volume transport on various cubed-sphere grids. *Journal of Computational Physics*, *227*(1), 55–78.

Quaas, J., Ming, Y., Menon, S., Takemura, T., Wang, M., Penner, J. E., … Schulz, M. (2009). Aerosol indirect effects – general circulation model intercomparison and evaluation with satellite data. *Atmos. Chem. Phys.*, *9*, 8697-8717. doi: 10.5194/acp-9-8697-2009

Rayner, N., Parker, D. E., Horton, E., Folland, C. K., Alexander, L. V., Rowell, D., … Kaplan, A. (2003). Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research: Atmospheres*, *108*(D14).

Ringer, M., Andrews, T., & Webb, M. (2014). Global-mean radiative feedbacks and forcing in atmosphere-only and coupled atmosphere–ocean climate change experiments. *Geophys. Res. Lett.*, *41*(11), 4035–4042.

Ringer, M., & et al. (2006). Global mean cloud feedbacks in idealized climate change experiments. *Geophys. Res. Lett.*, *33*(L07718).

Rosenblum, E., Fajber, R., Stroeve, J. C., Gille, S. T., Tremblay, L. B., & Carmack, E. C. (2021). Surface salinity under transitioning ice cover in the Canada basin: Climate model biases linked to vertical distribution of fresh water. *Geophys. Res. Lett.*, *48*.

Rotstayn, L., Ryan, B. F., & Katzfey, J. J. (2000). A scheme for calculation of the liquid fraction in mixed-phase stratiform clouds in large-scale models. *Mon. Wea. Rev.*, *128*, 1070-1088.

Rotstayn, L. D. (1997). A physically based scheme for the treatment of stratiform clouds and precipitation in large-scale models. I, Description and evaluation of the microphysical processes. *Q. J. R. Meteorol. Soc.*, *123*, 1227–1282.

Salzmann, M., Ming, Y., Golaz, J.-C., Ginoux, P. A., Morrison, H., Gettelman, A., … Donner, L. J. (2010). Two-moment bulk stratiform cloud microphysics in the GFDL AM3 GCM: description, evaluation, and sensitivity tests. *Atmospheric Chemistry and Physics*, *10*(16), 8037–8064. Retrieved from `https://www.atmos-chem-phys.net/10/8037/2010/` doi: 10.5194/acp-10-8037-2010

Sato, Y., Goto., D., Michibata, T., Suzuki, K., Takemura, T., Tomita, H., & Nakajima, T. (2018). Aerosol effects on cloud water amounts were successfully simulated by a global cloud-system resolving model. *Nature Communications*, *9*(985). doi: 10.1038/s41467-018-03379-6

Savijarvi, H. (1997). Shortwave optical properties of rain. *Tellus*, *49a*(2), 177–181. doi: 10.3402/tellusa.v49i2.14463

Schwarzkopf, D. M., & Fels, S. (1991). The simplified exchange method revisited: An accurate, rapid method for computation of infrared cooling rates and fluxes. *Journal of Geophysical Research*, *96*, 9075–9096.

Seifert, A., & Beheng, K. D. (2001). A double-moment parameterization for simulating autoconversion, accretion and selfcollection. *Atmospheric Research*, *59-60*, 265–281.

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*(e2019RG000678). doi: 10.1029/2019RG000678

Shevliakova, E., Pacala, S. W., Malyshev, S., Hurtt, G. C., Milly, P., Caspersen, J. P., … Crevoisier, C. (2009). Carbon cycling under 300 years of land use change: Importance of the secondary vegetation sink. *Global Biogeochemical Cycles*, *23*(2).

Soden, B. J., Broccoli, A., & Hemler, R. (2004). On the use of cloud forcing to estimate cloud feedback. *Journal of Climate*, *17*(19), 3661–3665. doi: doi:10.1175/1520-0442(2004)017⟨3661:OTUOCF⟩2.0.CO;2

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. doi: 10.1175/2007JCLI2110.1

Stephens, G., Li, J., Wild, M., Clayson, C. A., Loeb, N., Kato, S., … Andrews, T. (2012). An update on Earth's energy balance in light of the latest global observations. *Nature Geoscience*, *5*. doi: 10.1038/NGEO1580

Sui, C.-H., Li, X., & Yang, M.-J. (2007). On the definition of precipitation efficiency. *J. Atmos. Sci.*, *64*, 4506–4513. doi: doi.org/10.1175/2007JAS2332.1

Sui, C.-H., Li, X., Yang, M.-J., & Huang, H.-L. (2005). Estimation of oceanic precipitation efficiency in cloud models. *J. Atmos. Sci.*, *62*(12), 4358–4370. doi: doi.org/10.1175/JAS3587.1

Suzuki, K., Golaz, J. C., & Stephens, G. L. (2013). Evaluating cloud tuning in a climate model with satellite observations. *Geophysical Research Letters*, *40*(16), 4464–4468. doi: 10.1002/grl.50874

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

Tiedtke, M. (1993). Representation of clouds in large-scale models. *Mon. Wea. Rev.*, *121*(121), 3040–3061.

Toll, V., Christensen, M., Quaas, J., & Bellouin, N. (2019). Weak average liquid-cloud-water response to anthropogenic aerosols. *Nature*, *572*, 51–55. doi: 10.1038/s41586-019-1423-9

Torrence, C., & Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological society*, *79*(1), 61–78.

Voldoire, A., Saint-Martin, D., Sénési, S., Decharme, B., Alias, A., Chevallier, M., … et al. (2019). Evaluation of CMIP6 DECK experiments with CNRM-CM6-1. *J. Adv. Model. Earth Syst.*, *11*, 2177–2213. doi: 10.1029/2019MS001683

Vonder Haar, T. H., Bytheway, J. L., & Forsythe, J. M. (2012). Weather and climate analyses using improved global water vapor observations. *Geophys. Res.*

Lett., *39*(L16802). doi: 10.1029/2012GL052094

Wang, C., Soden, B. J., Yang, W., & Vecchi, G. A. (2021). Compensation between cloud feedback and aerosol-cloud interaction in CMIP6 models. *Geophys. Res. Lett.,*, *48*(e2020GL091024). doi: 10.1029/2020GL091024

Wang, M., Ghan, S., Liu, X., L'Ecuyer, T. S., Zhang, K., Morrison, H., … Penner, J. E. (2012). Constraining cloud lifetime effects of aerosols using A-Train satellite observations. *Geophys. Res. Lett.,*, *39*(L15709). doi: 10.1029/2012GL052204

Webb, M. J., Senior, C. A., Sexton, D. M. H., Ingram, W. J., Williams, K. D., Ringer, M. A., … et al. (2006). On the contribution of local feedback mechanisms to the range of climate sensitivity in two GCM ensembles. *Clim. Dyn.*, *27*, 17–38. doi: 10.1007/s00382-006-0111-2

Wheeler, M., & Kiladis, G. N. (1999). Convectively coupled equatorial waves: Analysis of clouds and temperature in the wavenumber-frequency domain. *Journal of the Atmospheric Sciences*, *56*, 374–399.

Wild, M., Folini, D., Schär, C., Loeb, N., Dutton, E. G., & König-Langlo, G. (2013). The global energy balance from a surface perspective. *Climate Dynamics*, *40*(11-12), 3107–3134. doi: 10.1007/s00382-012-1569-8

Winton, M., Adcroft, A., Dunne, J. P., Held, I. M., Shevliakova, E., Zhao, M., & et al. (2020). Climate sensitivity of GFDL's CM4.0. *Journal of Advances in Modeling Earth Systems*, *12*(e2019MS001838). doi: 10.1029/2019MS001838

Wittenberg, A. T. (2009). Are historical records sufficient to constrain ENSO simulations? *Geophysical Research Letters*, *36*, L14709. doi: 10.1029/2009GL038710

Wittenberg, A. T., Rosati, A., Delworth, T. L., Vecchi, G. A., & Zeng, F. (2014). ENSO modulation: Is it decadally predictable? *Journal of Climate*, *27*, 2667–2681. doi: 10.1175/JCLI-D-13-00577.1

Wittenberg, A. T., Vecchi, G. A., Delworth, T. L., Rosati, A., Anderson, W., Cooke, W. F., … Ray, S. (2018). Improved simulations of tropical Pacific annual-mean climate in the GFDL FLOR and HiFLOR coupled GCMs. *Journal of Advances in Modeling Earth Systems*, *10*, 3176–3220. doi: 10.1029/2018MS001372

Xiang, B., Harris, L., Delworth, T. L., B. Wang, G. C., Chen, J.-H., Clark, S. K., … et. al. (2021). S2S prediction in GFDL SPEAR: MJO diversity and teleconnections. *Bulletin of the American Meteorological Society*. doi: 10.1175/BAMS-D-21-0124.1

Yan, X., Zhang, R., & Knutson, T. R. (2018). Underestimated AMOC variability and implications for AMV and predictability in CMIP models. *Geophysical Research Letters*, *45*, 4319–4328. doi: 10.1029/2018GL077378

Yan, X., Zhang, R., & Knutson, T. R. (2019). A multivariate AMV index and associated discrepancies between observed and CMIP5 externally forced AMV. *Geophysical Research Letters*, *46*. doi: 10.1029/2019GL082787

Yano, J.-I., & Maarten, H. P. A. (2017). Moist static energy: definition, reference constants, a conservation law and effects on buoyancy. *Quart. J. Roy. Meteor. Soc.*, *143*, 2727–2734. doi: 10.1002/qj.3121

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*(e2019GL085782). doi: 10.1029/2019GL085782

Zhang, R., & Thomas, M. (2021). Horizontal circulation across density surfaces contributes substantially to the long-term mean northern Atlantic Meridional Overturning Circulation. *Communications Earth & Environment*, *2*(112). doi: 10.1038/s43247-021-00182-y

Zhao, J., & Johns, W. (2014). Wind-forced interannual variability of the Atlantic Meridional Overturning Circulation at $26.5^o N. Journal\ of\ Geophysical\ Research:\ Oceans, 119, 2403--2419.$

Zhao, M., Golaz, J.-C., Held, I., Guo, H., et al. (2018a). The GFDL global atmosphere and land model AM4. 0/LM4.0: 1. Simulation characteristics with prescribed SSTs. *Journal of Advances in Modeling Earth Systems*, *10*(3), 691–734.

Zhao, M., Golaz, J.-C., Held, I. M., Guo, H., et al. (2018b). The GFDL global atmosphere and land model AM4. 0/LM4.0: 2. Model description, sensitivity studies, and tuning strategies. *Journal of Advances in Modeling Earth Systems*, *10*(3), 735–769.

Zhao, M., Golaz, J.-C., Held, I. M., Ramaswamy, V., Lin, S.-J., Ming, Y., … Guo, H. (2016). Uncertainty in model climate sensitivity traced to representations of cumulus precipitation microphysics. *Journal of Climate*, *29*(2), 543–560. doi: 10.1175/JCLI-D-15-0191.1

Zhou, L., Lin, S.-J., Chen, J.-H., Harris, L., Chen, X., & Rees, S. L. (2019). Toward convective-scale prediction within the next generation global prediction system. *Bulletin of the American Meteorological Society*, *100*(7), 1225–1243. doi: 10.1175/BAMS-D-17-0246.1