The Meteorological Research Institute Earth System Model Version 2.0, MRI-ESM2.0: Description and Basic Evaluation of the Physical Component

Seiji YUKIMOTO, Hideaki KAWAI, Tsuyoshi KOSHIRO, Naga OSHIMA, Kohei YOSHIDA, Shogo URAKAWA, Hiroyuki TSUJINO, Makoto DEUSHI, Taichu TANAKA, Masahiro HOSAKA

Meteorological Research Institute, Tsukuba, Japan

Shokichi YABU

Japan Meteorological Agency, Tokyo, Japan

Hiromasa YOSHIMURA, Eiki SHINDO, Ryo MIZUTA, Atsushi OBATA

Meteorological Research Institute, Tsukuba, Japan

Yukimasa ADACHI

Japan Meteorological Agency, Tokyo, Japan

and

Masayoshi ISHII

Meteorological Research Institute, Tsukuba, Japan

(Manuscript received 27 November 2018, in final form 27 May 2019)

Abstract

The new Meteorological Research Institute Earth System Model version 2.0 (MRI-ESM2.0) has been developed based on previous models, MRI-CGCM3 and MRI-ESM1, which participated in the fifth phase of the Coupled Model Intercomparison Project (CMIP5). These models underwent numerous improvements meant for highly accurate climate reproducibility. This paper describes model formulation updates and evaluates basic performance of its physical components. The new model has nominal horizontal resolutions of 100 km for atmosphere and ocean components, similar to the previous models. The atmospheric vertical resolution is 80 layers, which is enhanced from the 48 layers of its predecessor. Accumulation of various improvements concerning clouds, such as a new stratocumulus cloud scheme, led to remarkable reduction in errors in shortwave, longwave, and net radiation at the top of the atmosphere. The resulting errors are sufficiently small compared with those in the CMIP5 models. The improved radiation distribution brings the accurate meridional heat transport required for the ocean and contributes to a reduced surface air temperature (SAT) bias. MRI-ESM2.0 displays realistic reproduction of both mean climate and interannual variability. For instance, the stratospheric quasi-biennial oscillation can now be realistically expressed through the enhanced vertical resolution and introduction of non-orographic gravity wave drag parameterization. For the historical experiment, MRI-ESM2.0 reasonably reproduces global SAT change.
Climate models are among the most effective tools for understanding the climate system and projecting possible future climate changes. Many climate models now include atmosphere chemistry and biogeochemistry processes, and thus have evolved into Earth system models. To improve climate change projection reliability for such Earth system models, we must first confirm that physical fields of the climate simulations are accurate and reliable before introducing complicated chemical processes to the model. For many years, climate models have been improved and have become more sophisticated; nevertheless, there are still uncertainties in many important physical aspects. For example, clouds play an essential role in Earth’s radiation budget and hydrological cycle, and although our understanding of clouds has deepened, there is still significant uncertainty (e.g., Boucher et al. 2013) and it remains challenging to represent clouds in climate models (Flato et al. 2013; Nam et al. 2012; Lauer and Hamilton 2013).

As demands for drafting climate change adaptation measures have increased in recent years, spatially and temporally detailed information on future climate changes is required. To respond to these demands, dynamical downscaling has been performed using regional climate models. Global climate models providing downscaling inputs are also required for precise regional climate simulations (Kitoh et al. 2016) in addition to high-fidelity representation of the global climate. Although representation of the summer monsoon in East Asia is improving (Sperber et al. 2013; Song and Zhou 2014), the monsoon rain belt known as the Baiu, which significantly influences Japanese water resources and industries, cannot be simulated with sufficient realism even by state-of-the-art global climate models (Kusunoki and Arakawa 2015).

The Meteorological Research Institute (MRI) of the Japan Meteorological Agency developed the MRI-CGCM3 global climate model (Yukimoto et al. 2012) and extended it to the MRI-ESM1 Earth system model (Yukimoto et al. 2011; Adachi et al. 2013), both of which contributed to the fifth phase of the Coupled Model Intercomparison Project (CMIP5; Taylor et al. 2012). Results from the CMIP5 experiments were open to climate research communities, leading to many studies that evaluated various aspects of model performance and contributed to the understanding of the climate system. MRI-CGCM3 has also been used for a broad range of studies such as climate change, paleoclimate, and climate extreme events (e.g., Kawai et al. 2016; Kaiho et al. 2016; Kaiho and Oshima 2017). The sixth phase of the CMIP (CMIP6), which is expected to fundamentally contribute to the Intergovernmental Panel on Climate Change’s Sixth Assessment Report, has been planned (Eyring et al. 2016). MRI has developed the new Earth system model MRI-ESM2.0 as a major update of MRI-ESM1, with many improvements to various model components such as vertical resolution enhancement of the atmospheric, oceanic, and chemical models. An emphasis on improving cloud and aerosol processes has been placed among others. With these improvements, MRI-ESM2.0 exhibits better performance than previous models in many aspects, and participating in CMIP6 with this model will contribute to activities in a wide range of research communities answering major scientific questions listed by the World Climate Research Program.

The purpose of this paper is to provide basic information of the MRI-ESM2.0 model performance derived from various model experiments and analyses of simulation outputs. We describe the formulations of MRI-ESM2.0 components, excluding those related to the carbon cycle, and describe the basic behavior of
2. Model description

MRI-ESM2.0 consists of four major component models: an atmospheric general circulation model (AGCM) with land processes, an ocean–sea-ice general circulation model, and aerosol and atmospheric chemistry models. MRI-ESM2.0 is an updated version of MRI-ESM1 (Yukimoto et al. 2011; Adachi et al. 2013) and has the same basic configuration as MRI-ESM1. In the remainder of this section, we describe model formulations with a focus on modifications from MRI-ESM1 (or its physical component MRI-CGCM3).

2.1. Atmosphere-land component

MRI-ESM2.0’s atmosphere-land component, MRI-AGCM3.5, is an updated version of MRI-AGCM3 adopted in MRI-CGCM3. MRI-AGCM3.5’s updates include 1) enhanced vertical resolution; improved parameterizations in 2) cloud macro- and micro-physics, 3) cloud radiation, and 4) gravity wave drag; and 5) modifications of aerosol microphysical and optical properties.

a. Resolution and vertical grid spacing

MRI-AGCM3.5 has a horizontal resolution of \( T_{159} \) (approximately 120 km), which was inherited from MRI-CGCM3. To prioritize doing many experiments with as large an ensemble size possible, we declined to enhance horizontal resolution in view of computational costs. In contrast, we enhanced atmospheric vertical resolution to better represent physical processes and atmospheric circulation in the model. The number of vertical levels increased from 48 to 80 (model top at 0.01 hPa) in a hybrid sigma-pressure (eta) coordinate system. By taking advantage of the semi-Lagrangian advection scheme (Yoshimura and Matsumura 2005; Yukimoto et al. 2011), the time step remains unchanged at 30 minutes, even if the vertical resolution is increased. Figure 1 shows the vertical grid spacing of the atmospheric models and related chemical models. Below 300 hPa, the grid spacing becomes finer than that of MRI-CGCM3. This grid spacing is the same as that of MRI-AGCM3.2, a high-resolution (with horizontal resolutions of \( T_{319} \) and \( T_{959} \)) atmospheric model used in a study of downscaling experiments for detailed future climate change over Japan (Mizuta et al. 2012). We decided to take advantage of MRI-AGCM3.2 as it had excellent precipitation field reproduction. A finer grid spacing is used above 300 hPa for better climate simulations in the upper troposphere, stratosphere, and mesosphere, including simulations of the stratospheric quasi-biennial oscillation (QBO; see Section 4.2 for further details). Well-resolved atmospheric waves and sharp structures are required, especially for better representations of the tropical tropopause layer, QBO, and hemispheric-scale meridional overturning circulation (i.e., Brewer–Dobson circulation). Grid spacing reaches a peak value smaller than 500 m at around 100 hPa, coarsens gradually upward. Compared to MRI-CGCM3, the grid spacing is approximately 1 km finer from the upper troposphere to the mesosphere. Note that in MRI-ESM2.0, background vertical diffusion is weakened in the stratosphere in accordance with finer vertical grid spacing.
b. Cloud macro- and micro-physics

Cloud macro- and micro-physics have been upgraded in several process formulations and parameters, which substantially improved both cloud representation and radiative flux.

A new stratocumulus scheme accounting for cloud top entrainment (Kawai et al. 2017; Kawai 2013) replaces the old stratocumulus scheme (Kawai and Inoue 2006). Because of this change, simulated low cloud cover over the subtropical oceans off the west coast of the continents and the Southern Ocean has significantly increased, leading to a faithful distribution of simulated low clouds.

Treatment of the Wegener–Bergeron–Findeisen (WBF) effect in cloud microphysics process has also been updated. The WBF effect is a deposition growth process of ice crystals at the expense of cloud droplets because ice saturation is lower than liquid water saturation. The WBF effect in MRI-CGCM3 was treated similarly to Lohmann et al. (2007). When the ice water content (IWC) is greater than 0.5 mg kg\(^{-1}\), the WBF effect in MRI-CGCM3 is treated similarly to Lohmann et al. (2007).

Because of this change, simulated low cloud cover over the subtropical oceans off the west coast of the continents and the Southern Ocean has significantly increased, leading to a faithful distribution of simulated low clouds.

Cloud droplet number concentration is estimated from aerosols in MRI-ESM2.0 and MRI-CGCM3. These models do not explicitly represent the Aitken mode of aerosols (see Table 1), although large enhancements of number concentration of sea salt aerosols, a part of which act as cloud condensation nuclei, are observed in this size range. To compensate for this effect, sea salt aerosol number concentration in the accumulation mode of the cloud process in MRI-

Table 1. Geometric mean radius \((r_m)\) (or number median radius) (μm) and standard deviation \((\sigma_r)\) in lognormal size distribution, hygroscopicity parameter \((\kappa)\), and complex refractive index at 550 nm for dry aerosols used in MRI-AGCM3.5 with references.

<table>
<thead>
<tr>
<th>Species</th>
<th>Size distribution ((r_m, \sigma_r))</th>
<th>Hygroscopicity parameter ((\kappa))</th>
<th>Complex refractive index at 550 nm (a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sulfate</td>
<td>((0.1, 1.80)) [SP06, L12]</td>
<td>0.53 [PK07]</td>
<td>(1.53–1.0 \times 10^{-7} ) i [T76]</td>
</tr>
<tr>
<td>Sulfuric acid(^d)</td>
<td>((0.3, 1.50)) [T18]</td>
<td>1.19 [PK07]</td>
<td>(1.43–1.0 \times 10^{-7} ) i [OPAC (SUSO)]</td>
</tr>
<tr>
<td>Hydrophobic BC(^e)</td>
<td>((0.0437, 1.64)) [S10]</td>
<td>(5 \times 10^{-7}) [G01]</td>
<td>(1.95–0.79) i [BB06, OPAC (SOOT)]</td>
</tr>
<tr>
<td>Hydrophilic BC(^e)</td>
<td>((0.0437, 1.64)) [S10]</td>
<td>0.53 [PK07]</td>
<td>(1.95–0.79) i [BB06, OPAC (SOOT)]</td>
</tr>
<tr>
<td>OM</td>
<td>((0.1, 1.80)) [SP06, L12]</td>
<td>0.12 [W08]</td>
<td>(1.53–6.0 \times 10^{-7} ) i [OPAC (WASO)]</td>
</tr>
<tr>
<td>Sea salt (accumulation mode)</td>
<td>((0.13, 2.03)) [SP06, OPAC]</td>
<td>1.12 [PK07]</td>
<td>(1.50–1.0 \times 10^{-8} ) i [OPAC (SSAM)]</td>
</tr>
<tr>
<td>Sea salt (course mode)</td>
<td>((1.75, 2.03)) [OPAC]</td>
<td>1.12 [PK07]</td>
<td>(1.50–1.0 \times 10^{-8} ) i [OPAC (SSCM)]</td>
</tr>
<tr>
<td>Mineral dust(^f)</td>
<td>6 size bins, ranging from 0.1 μm to 10 μm in particle radius</td>
<td>0.14 [G01]</td>
<td>(1.53–5.5 \times 10^{-3} ) i [OPAC (MINM, MIAM, MICM)]</td>
</tr>
</tbody>
</table>

\(^a\) SP06 = Seinfeld and Pandis (2006); L12 = Liu et al. (2012); T18 = Thomason et al. (2018); S10 = Schwarz et al. (2010); OPAC = Hess et al. (1998); PK07 = Petters and Kreidenweis (2007); G01 = Ghan et al. (2001); W08 = Wang et al. (2008); T76 = Toon et al. (1976); BB06 = Bond and Bergstrom (2006).

\(^b\) Letters in parentheses are file names used in OPAC. The complex refractive index of water is taken from OPAC (FOG).

\(^c\) Sulphate is assumed to be \((\text{NH}_4)_2\text{SO}_4\).

\(^d\) Sulfuric acid is assumed to be a mixture of 75 % \(\text{H}_2\text{SO}_4\) and 25 % water.

\(^e\) The complex refractive indices of black carbon (BC) in the visible range are taken from BB06 and those in the other range are taken from OPAC (SOOT). Hydrophobic BC is assumed to be dry without hygroscopic growth. Hydrophilic BC is assumed to be a mixture of BC and sulfate in the radiation process.

\(^f\) Size distributions of mineral dust are assumed to be lognormal for each bin range for calculations of aerosol optical properties in the look-up tables. The size distributions of OPAC (MINM), OPAC (MIAM), and OPAC (MICM) are used for bins 1, 2–3, and 4–6, respectively. Mineral dust is assumed to be dry without hygroscopic growth in the radiation process.
ESM2.0 is multiplied by a tuning factor of 2.0 based on observational studies (e.g., Covert et al. 1996; Clarke et al. 2006). In addition, the geometric mean radius of accumulation mode sea salt aerosols decreased from 0.228 μm (Chin et al. 2002) to 0.13 μm (Seinfeld and Pandis 2006), thus increasing cloud droplet number concentration originating from sea salt aerosols. These modifications increase cloud optical depth over the ocean, including the Southern Ocean.

Additionally, several cloud scheme defects, including an incorrect treatment associated with the prognostic equations of cloud particle number concentrations (Kawai et al. 2015; Tsushima et al. 2016) have been fixed in MRI-ESM2.0. These modifications and other minor changes to cloud schemes and their impacts on cloud representation in MRI-ESM2.0 are discussed in detail in Kawai et al. (2019).

c. Cloud radiation

The cloud overlap assumption for shortwave radiation was updated. For longwave radiation calculation, a maximum-random overlap is assumed as a vertical arrangement of clouds (Geleyn and Hollingsworth 1979) in both MRI-CGCM3 and MRI-ESM2.0. In contrast, for shortwave radiation in MRI-CGCM3, total cloud cover in a column was first computed based on the maximum-random overlap assumption, and then random overlap was assumed to solve radiative fluxes in a cloudy sub-column. This procedure was chosen for computing efficiency; however, the treatment led to excessive reflection of solar radiation, especially for deep convection towers topped by optically thin anvil clouds (Nagasawa 2012). In MRI-ESM2.0, the maximum-random overlap assumption becomes usable for calculating shortwave radiation because of the introduction of a practical independent column approximation (Nagasawa 2012) based on Collins (2001). The new cloud overlap scheme drastically reduces excessive shortwave radiation reflection over tropical convection areas.

d. Gravity wave drag

In MRI-ESM2.0, subgrid-scale forcing due to orographic gravity wave drag (OGWD) and non-orographic gravity wave drag (NGWD) is parameterized. MRI-ESM2.0’s OGWD scheme of Iwasaki et al. (1989) is inherited from MRI-CGCM3. To better represent middle atmosphere circulation, the NGWD scheme of Hines (1997) was newly introduced along with vertical resolution enhancement. Except for some tunable parameters, the NGWD scheme setup is the same as that used in Shibata and Deushi (2005), which simulated realistic QBO-like wind variations. NGWD’s isotropic source strength (assuming root mean square wind perturbations) is about 2.0 m s$^{-1}$ everywhere. The isotropic source launch height is set to approximately 700 hPa. Above the upper stratosphere, Rayleigh friction, which was introduced as a proxy of gravity wave drag in MRI-CGCM3, is weakly imposed to fill up the lack of the gravity wave drag near the model top, with relaxation times of approximately 30, 5, and 3 days at around 1, 0.1, and 0.01 hPa, respectively. Introducing NGWD, as along with fine vertical grid spacing, plays a key role in simulating stratospheric QBO (e.g., Baldwin et al. 2001; Watanabe et al. 2008; Kawatani et al. 2010) as shown in Section 4.2j.

e. Aerosol microphysical and optical properties

In MRI-ESM2.0, aerosol quantities used in cloud microphysics and aerosol-radiation calculations in the atmospheric component MRI-AGCM3.5 are provided from the coupled aerosol model (see Section 2b). Aerosol mass concentrations provided are converted to number concentrations with prescribed aerosol size distributions in MRI-AGCM3.5 (Table 1). MRI-ESM2.0 separately treats volcanic origin sulfuric acid aerosol from sulfate aerosol, and it also separately treats hydrophobic and hydrophilic black carbon (BC) aerosols. The atmospheric component of MRI-CGCM3 treated the former and the latter as total sulfate and BC aerosols, respectively. For more reliable estimates of aerosol effects on radiation and clouds in MRI-ESM2.0, we updated parameters of lognormal size distribution and complex refractive index of aerosols and introduced a hygroscopicity parameter, κ, in MRI-AGCM3.5, based on recent observations and experiments (Table 1). In MRI-CGCM3, aerosol size distribution parameters were taken from Chin et al. (2002) and the Optical Properties of Aerosols and Clouds (OPAC) database (Hess et al. 1998). However, Chin et al. (2002) and OPAC have smaller size-distribution parameters of fine aerosol particles than observed values (e.g., Takegawa et al. 2009), which has likely led to large errors in radiative effect estimates in MRI-CGCM3. In MRI-ESM2.0, we adopted size-distribution parameters for BC based on airborne measurements, and used values reported in recent literature for sulfate and organic matter (OM), assuming internal mixing (Table 1). The lognormal size distribution of sulfuric acid aerosol is estimated from volume density, surface area density, and mean radius of stratospheric aerosols in the stratospheric aerosol dataset used in the CMIP6 experiments (Thomason et al.
The hygroscopic growth factors of particle radius with ambient relative humidity in MRI-CGCM3 were taken from Chin et al. (2002) and OPAC, and were separately used in optical and other processes, respectively, in an inconsistent manner. MRI-ESM2.0 theoretically calculates the hygroscopic growth factors of each chemical component by Köhler theory using the corresponding hygroscopicity parameter, $\kappa$, (Table 1) which represents a quantitative measure of aerosol water uptake characteristics and cloud condensation nuclei activity (Petters and Kreidenweis 2007). Radiation and cloud processes use common hygroscopicity parameters (e.g., Abdul-Razzak and Ghan 2000). In MRI-CGCM3, the spectra complex refractive indices were taken from the OPAC database (Hess et al. 1998). We updated these values for BC in a visible range and sulfate in MRI-ESM2.0 because previous studies demonstrated weak absorptive properties in OPAC values (e.g., Bond and Bergstrom 2006) and because OPAC values for stratospheric sulfate droplets (mixture of 75% H$_2$SO$_4$ and 25% water) were used for tropospheric sulfate in MRI-CGCM3.

Normalized optical properties of aerosols (i.e., extinction coefficient, single scattering albedo, and asymmetry factor given by per unit aerosol number concentration for each aerosol category) in the solar and terrestrial spectral range (i.e., nine longwave and 22 shortwave bands) were pre-calculated as look-up tables based on Mie theory, assuming spherical particles using aerosol microphysical parameters and complex refractive indices (Table 1) in a consistent manner. In MRI-CGCM3’s radiation process, all BC was treated as hydrophobic (i.e., BC had neither hygroscopic growth nor coatings), which likely led to a significant underestimation of light absorption by BC. The optical properties of hydrophilic BC in MRI-ESM2.0 are calculated based on Mie theory with a core-shell aerosol treatment, in which a concentric core of BC is surrounded by a uniform coating shell of other aerosol compounds (Oshima et al. 2009a, b). We assume internal mixing of BC and sulfate with a shell-to-core volume ratio of 2, which was obtained by airborne measurements (Moteki et al. 2012, 2017), and assume the same hygroscopic growth as coating species (i.e., sulfate) for hydrophilic BC. The complex refractive indices for other hygroscopic aerosol species are calculated by volume-weighted averaging of refractive indices of aerosol components and water. Finally, the optical properties of aerosols (e.g., aerosol optical depths and scattering and absorption coefficients) are calculated using the look-up tables with ambient relative humidity and aerosol number concentrations.

MRI-ESM2.0’s radiative transfer process follows a method used in MRI-CGCM3. The direct radiative effect of aerosols was extremely weak in MRI-CGCM3 (e.g., Cherian et al. 2014; Zelinka et al. 2014) because the aerosol number concentrations in the radiation scheme were treated incorrectly; however, this has been corrected in MRI-ESM2.0.

### f. Other atmospheric sub-components

The cumulus convection scheme (Yoshimura et al. 2015) is basically unchanged; however, a modification was added to suppress shallow convection when a stratocumulus cloud is diagnosed. Additionally, some parameters were tuned. For example, the precipitation conversion rate in the updraft was slightly increased so cloud water detrained less at the middle levels. Full radiation calculations are performed for every grid in MRI-ESM2.0, whereas in MRI-CGCM3 they were performed for every two grids to reduce computational costs. Basically, there are no changes from MRI-CGCM3 (Yukimoto et al. 2012) in the other atmospheric sub-components of land surface model, radiation code, and boundary layer scheme.

### 2.2 Aerosol model

Model of Aerosol Species in the Global Atmosphere mark-2 revision 4-climate (MASINGAR mk-2r4c) is an aerosol model that is a component of MRI-ESM2.0, and is an updated version of MASINGAR mk-2 (Tanaka et al. 2003; Tanaka and Chiba 2005; Yumimoto et al. 2017; Tanaka and Ogi 2017) that was used in MRI-CGCM3. Because a basic structure and formulation of MASINGAR mk-2r4c in MRI-ESM2.0 follows those of MASINGAR mk-2 in MRI-CGCM3, a brief description of the model and major modifications from MRI-CGCM3 is given below. Model modification details will be described elsewhere (Oshima et al. in preparation).

MASINGAR mk-2r4c treats atmospheric aerosol physical and chemical processes (e.g., emission, transport, diffusion, chemical reactions, and dry and wet depositions) and includes the following species: nonsea-salt sulfate, BC, organic carbon (OC), sea salt, mineral dust, and aerosol precursor gases (e.g., sulfur dioxide [SO$_2$] and dimethyl sulfide). This model assumes external mixing for all aerosol species; however, in MRI-AGCM3.5’s radiation process, hydrophilic BC is assumed to be internally mixed with sulfate. In the CMIP6 experiments, sulfur species originating from anthropogenic, biogenic, and volcanic sources are treated separately. In MASINGAR
2.3 Atmospheric chemistry

MRI-ESM2.0’s chemistry component is the MRI Chemistry Climate Model version 2.1 (MRI-CCM2.1), which simulates the distribution and evolution of ozone and other trace gases in the troposphere and middle atmosphere. In the simulations, meteorological fields and aerosol surface area density are received from the interactively coupled MRI-AGCM3.5 and MASINGAR mk-2r4c models. For the CMIP6 experiments, MRI-CCM2.1 employs a T42 horizontal resolution (approximately 280 km, which is different from that of MRI-AGCM3.5) and 80 vertical layers (same as MRI-AGCM3.5).

MRI-CCM2.1 is an updated version of MRI-CCM2 (Deushi and Shibata 2011), which was used as MRI-ESM1’s atmospheric chemistry component. To improve ozone chemistry processes in the upper stratosphere and mesosphere, MRI-CCM2.1 includes several important updates. First, it includes 12 chemical reactions (Table 2) that were not in the previous version. MRI-CCM2.1 calculates a total of 90 chemical species and 259 chemical reactions. Second, we updated kinetic reaction rates from the Jet Propulsion Laboratory database compiled by Sander et al. (2006; the JPL06-2 database) to those from the JPL10-6 database (Sander et al. 2011). Third, volume mixing ratios of various short-lived species are diagnosed at each time step with a similar framework to the daytime chemical reaction scheme. These species include OH, Cl, O(3P), HO2, Br, N, H, C2H2O2, C2H4O2, ACETO2, EO2, EO, and PO2. In the previous version, the mixing ratios were set to a very small constant value (1 × 10−36) during nighttime, which led to significant errors in simulated ozone abundance, especially in the upper stratosphere and the mesosphere. Fourth, the effects of energetic particle precipitation (EPP) on atmospheric
ozone chemistry are newly introduced for the CMIP6 experiments following Matthes et al. (2017). The EPP odd nitrogen (NOy) upper boundary condition, in which NOy (= NOx) molecular fluxes are forced at the model top (0.01 hPa), is implemented in MRI-CCM2.1 to account for the EPP indirect effect coming from above the model top. Additionally, NOy and HOy production due to particle-induced ionization in the model domain (i.e., the EPP direct effect) is also implemented following Appendices D and E of Matthes et al. (2017).

2.4 Ocean and sea ice

MRI-ESM2.0’s ocean–sea-ice component is the MRI Community Ocean Model version 4 (MRI.COMv4; Tsujino et al. 2017), which is an upgraded version of the MRI.COM3 (Tsujino et al. 2010) model used in MRI-CGCM3. MRI.COMv4 is a free-surface depth-coordinate ocean–sea-ice model that solves primitive equations using Boussinesq and hydrostatic approximations on a structured mesh. Here, we briefly describe the ocean component settings, focusing on updates from the previous model. The MRI.COMv4 reference manual (Tsujino et al. 2017) describes settings in detail.

MRI.COMv4 uses a horizontal grid system that is almost identical to MRI.COM3, adopting Murray’s (1996) tripolar grid with a nominal horizontal resolution of 1-degree in longitude and 0.5-degrees in latitude. However, the meridional resolution between 10°S and 10°N was increased to 0.3 degrees, whereas the meridional resolution was a uniform 0.5° in MRI.COM3. The vertical grid system uses a vertically rescaled height coordinate (z* coordinate) proposed by Adcroft and Campin (2004), which allows shallower (8-m minimum) sea-floor depths than the previous model (32-m minimum). The number of vertical layers increased from 50 to 60 and the nominal thicknesses range from 2 m in the top layer to 700 m in the bottom layer. To better resolve the mixed layer and seasonal thermocline, layer thicknesses do not exceed 10 m in the upper 200 m. The updated model also has a bottom boundary layer from Nakano and Sugino-Hara (2002), which has a thickness of 50 m, as was the case in the previous model. Topography is based on the General Bathymetric Chart of the Oceans 2014 version 20141103 (Weatherall et al. 2015) merged with the JTOPO30v2 depth data product for the western North Pacific region published by the Marine Information Research Center.

MRI-ESM2.0 continually adopts the generalized Arakawa scheme as described by Ishizaki and Motoi (1999) for momentum advection terms and uses the second order moment scheme of Prather (1986) for tracer advection terms (potential temperature, salinity, and ideal age tracer), like the old model. In the updated model, the flux limiter for Prather (1986)’s scheme was replaced with method B proposed by Morales Maqueda and Holloway (2006).

Several modifications have been introduced to sub-grid scale parameterizations in MRI-COMv4. We continually use a flow-dependent anisotropic horizontal viscosity scheme from Smith and McWilliams (2003), but the anisotropic factor for viscosity perpendicular to a local flow linearly tapers from 0.2 at 5°N/S to 0.1 at the equator in the present model to improve flow speed of the Equatorial Undercurrent. Additionally, we introduced several updates to the vertical mixing. The present model adopts the generic length scale scheme of Umlauf and Burchard (2003) as a turbulence closure scheme for vertical mixing. Although the previous scheme (Noh and Kim 1999) only solves the prognostic equation of turbulent kinetic energy, the new scheme also solves the other prognostic equation of the (generic) length scale. Background vertical diffusion coefficients have a three-dimensional empirical distribution based on Decloedt and Luther (2010), whereas in the previous version they were horizontal constant. Recent observational studies have revealed that turbulent vertical diffusivity substantially varies in space (e.g., Polzin et al. 1997; Ledwell et al. 2000; Hibiya et al. 2006). To incorporate this feature into our new model, we adopted the three-dimensional distribution mentioned above and stopped using the one-dimensional vertical profile of Tsujino et al. (2000). Vertical diffusivity was locally set to a large value of 1 m² s⁻¹ whenever unstable stratification is detected in the model.

Both the present and previous models use the isopycnal tracer diffusion scheme of Redi (1982) and the eddy-induced tracer transport scheme of Gent and McWilliams (1990) to parameterize stirring due to mesoscale eddies; however, the coefficients of these schemes were extensively modified in the present model. The constant isopycnal tracer diffusivity was enlarged from 1000 m² s⁻¹ in MRI-CGCM3 to 1500 m² s⁻¹ in MRI-ESM2.0. We introduced two tapering factors for the oceanic interior and a layer near the sea surface as proposed by Danabasoglu and McWilliams (1995; their Eq. A.7a with S_c = 0.08 and S_d = 0.01) and Large et al. (1997; their Eq. B.4), respectively. These tapering factors were applied to all but the horizontal-diagonal elements of the isopycnal tracer diffusion tensor. Therefore, the isopycnal diffusion
is gradually modified to horizontal diffusion around steeply tilted isopycnal surfaces and within the surface diabatic layer. The coefficient for Gent and McWilliams parameterization is calculated with schemes of Danabasoglu and Marshall (2007) and Danabasoglu et al. (2008). The coefficient depends on a local buoyancy frequency and ranges from 300 m$^2$ s$^{-1}$ in weakly stratified regions to 1500 m$^2$ s$^{-1}$ in strongly stratified regions. Within the surface diabatic layer, parameterized eddy-induced transport is modified, so that a corresponding meridional overturning streamfunction linearly tapers to zero from the bottom of the diabatic layer to the sea surface. We put a ceiling at 0.005 on the isopycnal slope evaluated in this parameterization.

The sea-ice component of MRI-ESM2.0 is almost unchanged from MRI-CGCM3. Thermodynamics are based on Mellor and Kantha (1989). Categorization of thickness, ridging, and rheology are adopted from the Los Alamos National Laboratory sea-ice model (Hunke and Lipscomb 2006). Fractional area, snow volume, ice volume, ice and surface temperature of each thickness category are transported using the multidimensional positive definite advection transport algorithm of Smolarkiewicz (1984). The current model also includes some bug fixes and minor changes in settings.

### 2.5 Coupling

Each component model is coupled by a coupler, known as Scup (Yoshimura and Yukimoto 2008), that is incorporated in MRI-ESM2.0 as well as in MRI-CGCM3 and MRI-ESM1 as an interface for coupling component models with different grids. Scup interpolates model outputs to the receiving component’s grid following a prescribed rule based on which component sends (receives) variables to (from) the other component at specified time intervals. In the CMIP6 experiments, MRI-ESM2.0 has a coupling interval set to 1 hour for all variables. In the case of coupling between the atmospheric and ocean–sea-ice components, the sea-ice component sends surface and interior temperatures, snow and ice thicknesses, and sea-ice fractional area to the atmospheric component model, and receives surface fluxes calculated by the atmospheric component model via Scup. The global-mean amount is strictly preserved in surface flux interpolations. For coupling involving the atmospheric component, the aerosol and atmospheric chemistry models, three-dimensional atmospheric variables (e.g., meteorological fields, aerosols, and chemical compositions) are also exchanged along with two-dimensional surface properties.

#### 2.6 Tuning

First, tuning was done for each component model, and then coupled tuning was done. In the tuning of the atmospheric model, we adjusted a combination of parameters so that errors in the annual mean top of atmosphere (TOA) radiation distribution become sufficiently small in 5-year test runs with present-day sea-surface conditions. The tuning parameters include precipitation conversion rate (‘rcn_akdz_rate’, ‘rcn_no_prec_height’) and maximum turbulent entrainment–detrainment rate (‘rcn_entdet_rate_2’) in the cumulus convection scheme, mode radius (‘rm_seasalt_bin1’) and a number concentration tuning factor (‘fac_nssalt_finer’) of fine-mode sea-salt aerosol, threshold parameters for diagnosing stratocumulus clouds (‘eishrs’, ‘critctei’), and a parameter for ice crystals to snow conversion (‘raggregat’). In doing the atmospheric tuning, it was sometimes necessary to examine climate sensitivity to ensure that the climate sensitivity did not fall below that of MRI-CGCM3, in which the 20th century temperature rise was underestimated (as shown in Section 4.3). This means that we weakly constrained the model’s temperature rise in the 20th century. We experienced that ‘rcn_entdet_rate_2’ has a moderate sensitivity to the climate sensitivity. As reported in Golaz et al. (2013), we found that the aerosol’s forcing in our model also greatly changes in response to the effective radius threshold for cloud-precipitation conversion. However, we did not change the threshold.

In the tuning after coupling, we slightly modified ‘critctei’ and ‘fac_nssalt_finer’, which are insensitive to the climate sensitivity and the aerosol’s radiative forcing, for tuning the global-mean TOA net radiation and reducing model drift under preindustrial conditions. In a preliminary historical experiment, the model simulated too large a variability in the equatorial Pacific. Since reducing ‘rcn_entdet_rate_2’ was found to have an effect to suppress the variability, we selected a slightly lower value for the parameter (that slightly increases the climate sensitivity) and re-tuned for the preindustrial spin-up by adjusting ‘critctei’ and ‘fac_nssalt_finer’.

### 3. Experimental design

In CMIP6 (Eyring et al. 2016), the preindustrial control (piControl) experiment, one of the Diagnostic, Evaluation, and Characterization of Klima experiments, and the historical experiment are used to establish basic model characteristics. In this section,
we describe these experiments’ designs and how the model climate system’s initial state was made.

3.1 Spin-up procedure

Before coupling the atmosphere and ocean, we took initial atmospheric and land surface data from ERA-Interim for 00 UTC, January 1, 2009. The initial oceanic state data were taken from a 342-year ocean–sea-ice model integration forced by the interannually varying atmospheric boundary conditions of the Japanese 55-year reanalysis (JRA-55; Kobayashi et al. 2015) converted as the surface dataset of version 1.1 for the Ocean Model Intercomparison Project (JRA-55-do; Tsujino et al. 2018). Initial distributions of aerosol species and their precursor gases were set to zero, except for carbonyl sulfide below 100 hPa, which was set to 500 pptv. Present-day climatological states obtained from Deushi and Shibata (2011) were used as initial data on ozone and other trace gases.

From these initial states, a spin-up run of MRI-ESM2.0 was performed for 1000 years under piControl conditions (see below) to obtain the initial state for the piControl experiment. Some minor modifications were introduced to the model during this spin-up phase. At year 501, a negligible bug in the grid conversion for surface wind stress received by the ocean–sea-ice model from the atmospheric model was fixed. At year 651, the format of NO\textsubscript{3} boundary conditions for the atmospheric chemistry component was changed from concentration to flux. At year 851, the sea-ice strength calculation was corrected for the ocean–sea-ice model. These modifications did not significantly affect the spin-up state. Although the official CMIP6 forcing dataset version 6.0 was used for the first 650 years of the 1000-year spin-up run, the dataset was updated several times. We used updated datasets in integrations version 6.1.1 from year 651 and version 6.2.1 from year 801.

3.2 piControl experiment

The piControl experiment serves as a baseline for evaluating the state of coupled climate system. Additionally, various CMIP6 simulations branch from the piControl experiment and are thus compared with this. The piControl experiment, with a 500-year run, started from the final spin-up state at year 1000. Official CMIP6 forcing dataset version 6.2.1 was used here. Solar forcing and stratospheric aerosols in the piControl experiment were given, which were monthly climatologies constructed from historical data (see Section 3.3) for 1850–1873 for solar forcing and 1850–2014 for stratospheric aerosols. All other external forcing agents were set to the same values as those at year 1850 in the historical experiment (see Section 3.3 for details). Concentrations of CO\textsubscript{2}, CH\textsubscript{4}, and N\textsubscript{2}O were set to 284.32 ppmv, 808.25 ppbv, and 273.02 ppbv, respectively.

3.3 Historical experiment

The CMIP6 historical experiment is designed to assess the model’s ability to simulate the present climate and climate changes of the recent past. The historical period in question is 1850–2014. The official CMIP6 forcing dataset version 6.2.1 was also used in the historical experiment. Historical records of greenhouse gas concentrations, emissions of anthropogenic short-lived climate forcers, open biomass burning emissions, and solar radiative and particle forcings were compiled by Meinshausen et al. (2017), Hoesly et al. (2018), van Marle et al. (2017), and Matthes et al. (2017), respectively.

In MRI-ESM2.0, volcanic aerosols are given by the conjunction of the stratospheric aerosol dataset used in the CMIP6 experiments (Thomason et al. 2018) and interactive aerosol model calculations. Mass concentration of sulfuric acid aerosol for radiation calculation is derived from aerosol volume density (assuming a mixture of 75 % H\textsubscript{2}SO\textsubscript{4} and 25 % water) in the stratospheric (above the climatological tropopause) aerosol dataset. For volcanic forcing originating from continuous volcanic eruptions, we used sulfuric acid aerosol concentrations interactively calculated by MASINGAR mk-2r4c, which uses the Global Emissions Inventory Activity database SO\textsubscript{2} emission inventory (Andres and Kasgnoc 1998). The latter concentrations are used only if they are greater than the former concentrations that primarily reflect sporadic volcanic eruptions.

Historical land-use changes were evaluated using the Land-Use Harmonization (LUH2) dataset provided by the CMIP6 land-use group (Lawrence et al. 2016). A forest area ratio function is calculated as the index of change of the vegetation type from forest to grass for each model grid. By using this index, past land-use changes are defined relative to the standard vegetation type, which is assumed to be the 1990 vegetation distribution.

Ensemble simulations are useful for separating forced climate signals from internal climate variations. We performed five ensemble simulations of the historical experiment. Each initial state was taken from January 1 of years 1 (i.e., the final state of the 1000-year spin-up run), 51, 101, 151, and 201 of the piControl experiment.
4. Model results

In this section, we describe climate drift in MRI-ESM2.0’s spin-up and piControl simulations, and then evaluate present-day climate and historical climate change in the historical experiment compared with the observations and results from MRI-CGCM3. Because MRI-ESM1 and MRI-CGCM3 have common physical components and reproduce similar climates with each other, we will compare the results of MRI-ESM2.0 with those of MRI-CGCM3 only. Unless otherwise noted, the present-day climate is a five-member ensemble-mean for 1986–2005 average in the historical experiment for both MRI-CGCM3 and MRI-ESM2.0. Note that the MRI-CGCM3 historical experiment is driven by forcing of the CMIP5 protocol and is executed for 1850–2005 (Yukimoto et al. 2012).

4.1 Climate drift

A coupled climate model used for climate change projection and climate sensitivity studies should have its control simulation with a stable model climate on a temporal scale greater than 100 years at least near the surface. Moreover, it is desirable that the model be in equilibrium in terms of global energy and water budgets. Figure 2 shows time evolutions of global averages of important climate variables over 1500 years (1000-year spin-up run plus 500-year piControl experiment). Surface air temperature (SAT, Fig. 2a) decreased steeply in the first few decades of the spin-up period because of the preindustrial radiative forcing given to the present-day initial state. After that, SAT increased slowly for the rest of the simulation. The rate of increase gradually decreased and the temperature rose by 0.5°C within the spin-up period.

![Figure 2](attachment:image.png)

Fig. 2. Time evolution of global averages of (a) surface air temperature (SAT) (°C), (b) sea-surface temperature (SST) (°C), (c) radiation budget (positive downward) at the top of the atmosphere (TOA) (W m⁻²), (d) energy budget at the ocean surface including energy accompanying river runoff and iceberg discharge (W m⁻²), (e) volume-averaged ocean temperature (°C), and (f) sea-surface salinity (SSS) (psu) of MRI-ESM2.0 over the spin-up and piControl simulations. The vertical dashed line indicates the piControl experiment start year. Values in the upper right corner are the average and trend per century for the piControl experiment period (500 years).
An increase rate of 0.017°C/century remains in the piControl experiment. The evolution of the sea-surface temperature (SST, Fig. 2b) is similar to that of SAT except that the short-term variation and the long-term change are both relatively small. SST trend in the piControl experiment period is 0.015°C/century.

The TOA radiation budget (Fig. 2c) and ocean surface energy input (Fig. 2d) show similar temporal evolutions. Both increased during the first few spin-up decades and then decreased slowly toward the end of the piControl experiment period. The piControl experiment has average energy inputs of 0.94 W m$^{-2}$ at the TOA and 0.22 W m$^{-2}$ at the ocean surface (0.16 W m$^{-2}$ for the global area). The majority of the energy budget discrepancy between the TOA and the ocean surface is due to known problems. About half of the discrepancy is explained by the internal energy of water vapor and precipitation ignored in the atmospheric model. And about one quarter of this portion can be explained by improper treatment of rain falling on sea ice as snow.

It has been confirmed that this artificial energy sink is mostly unchanging over time and has little effect on climate change and climate sensitivity in preliminary sensitivity experiments. For example, when the carbon dioxide concentration abruptly quadrupled, the energy sink varied only 0.17 W m$^{-2}$/century, which is sufficiently small compared to TOA radiation changes greater than 6 W m$^{-2}$. Consistent with excess energy input to the ocean, the volume-averaged ocean temperature (Fig. 2e) continues to increase with an average trend of 0.032°C/century in the piControl experiment period. This indicates that although the sea surface is relatively thermally stable, the deep ocean has not reached thermal equilibrium.

The sea-surface salinity (SSS, Fig. 2f) increased relatively rapidly during the first half of the spin-up period. After that, the increase rate decreases over time, reaching almost zero during the piControl experiment. This implies that, at least on Earth’s surface, the hydrological cycle is almost in equilibrium.

### 4.2 Present-day climate

#### a. Radiation

Radiation at the TOA is the most fundamental factor for a climate model, and thus must be verified against reliable observational data. To simulate radiation distributions at the TOA as accurately as possible, we focused model development on clouds, with expressions of cloud distribution, including cloud overlap in the radiation scheme and cloud properties such as cloud-aerosol interactions, as described in the previous section.

Figure 3 shows the biases of annual-mean shortwave, longwave, and net radiation at the TOA in the present-day climate against CERES-EBAF Edition 2 (Loeb et al. 2009; 2001–2010 average). In MRI-CGCM3, there was a negative bias for shortwave radiation, that is, excessive reflection generally in the tropics; however, in the middle and high latitudes, particularly in the Southern Ocean, there was a large positive bias. These biases were found to be dramatically improved in MRI-ESM2.0. The global average bias of MRI-ESM2.0 was $+0.21$ W m$^{-2}$, far lower than the $−1.95$ W m$^{-2}$ of MRI-CGCM3, and the root mean square error (RMSE) also significantly decreased, going from 16.84 W m$^{-2}$ in the previous model to 9.85 W m$^{-2}$. Here and hereafter, RMSE is calculated without subtracting the global mean. Excessive reflections in the tropics are reduced, which can mostly be attributed to improvements related to cloud overlap in the radiation scheme. At mid-to-high latitudes, particularly over the Southern Ocean, the overestimated solar absorption characteristic of MRI-CGCM3 has almost disappeared in MRI-ESM2.0. Excessive solar absorption in the Southern Ocean was a lingering problem many climate models and reanalyses had faced for many years (e.g., Trenberth and Fasullo 2010). Although the cause of this issue was not fully understood, it was recently indicated that expressions of middle and low clouds in cold sectors of cyclones are insufficient (Bodas-Salcedo et al. 2012). Additionally, supercooled cloud water, which has been found to be ubiquitous in those clouds, is not well simulated (Cesana and Chepfer 2013). We believe that introducing a new stratocumulus cloud parameterization applicable to mid-to-high latitudes and improvements of cloud microphysics have contributed to improving this bias in MRI-ESM2.0.

Patterns for longwave radiation are similar between MRI-CGCM3 and MRI-ESM2.0; however, bias magnitude for MRI-ESM2.0 is generally much smaller. MRI-ESM2.0 shows remarkable improvements over MRI-CGCM3 for global-mean bias ($+0.70$ versus $+2.34$ W m$^{-2}$) and RMSE (6.85 versus 10.44 W m$^{-2}$). For net radiation, large errors in shortwave and longwave radiation compensated each other in MRI-CGCM3. In contrast, MRI-ESM2.0 has much smaller errors for shortwave and longwave radiation than MRI-CGCM3, resulting in small net radiation errors; however, a similar but weaker compensation of radiation errors can still be seen in the tropics.

We evaluated how MRI-ESM2.0’s TOA radiation distribution accuracy is superior to that of the CMIP5 multi-models, including MRI-CGCM3 (Fig. 4). For a
Fig. 3. Biases of the annual-mean radiation at the top of the atmosphere (TOA) relative to the CERES-EBAF observations for MRI-CGCM3 (left) and MRI-ESM2.0 (right), for shortwave (upper), longwave (middle), and net (bottom) radiations. Unit is W m\(^{-2}\) (positive downward).

Fig. 4. Taylor diagram for the annual-mean (a) shortwave, (b) longwave, and (c) net radiations at the top of the atmosphere (TOA). The reference point on the horizontal axis at which the normalized standard deviation is 1.0 corresponds to the CERES-EBAF observations. Plots are shown for the MRI-CGCM3 (blue dot), MRI-ESM2.0 (red dot), the CMIP5 multi-model mean (black square), and individual CMIP5 models (crosses).
fair comparison, we used only one ensemble member for each model. In Taylor diagrams for shortwave, longwave, and net radiation, MRI-CGCM3 was near average or in places representing relatively poor performance among the CMIP5 models. In contrast, MRI-ESM2.0 appears to be the best of the CMIP5 models, with a performance comparable to that of the CMIP5 multi-model mean.

b. Surface air temperature

SAT significantly influences human activity; thus, it is an important variable for assessing climate change impacts. Moreover, because SAT is strongly affected by radiation distribution and other climate factors, it is also one of the most fundamental variables to accurately simulate in climate models. Figure 5 illustrates bias distributions of present-day climate annual-mean SAT (1986–2005 average) reproduced in the historical experiments. The biases are relative to JRA-55 (Kobayashi et al. 2015). MRI-CGCM3 and MRI-ESM2.0 both represent the overall SAT distribution well, with a spatial correlation greater than 0.99 against the reanalysis. However, as a whole, MRI-ESM2.0 has more accurately simulated SAT than MRI-CGCM3. MRI-ESM2.0 has smaller bias (−0.23°C) and RMSE (1.31°C) than MRI-CGCM3 (−0.60°C and 1.93°C, respectively). There was a large-scale bias distribution in MRI-CGCM3 (Fig. 5a), with prominent and extensive cold and warm biases in the Northern and Southern hemispheres, respectively. In contrast, MRI-ESM2.0 (Fig. 5b) showed no biases exceeding 2°C except for at the poles, parts of the central North Atlantic, the central North Pacific, and other small areas. The North Atlantic in particular is almost covered by a cold bias exceeding −2°C in MRI-CGCM3, but this cold bias is greatly reduced in MRI-ESM2.0. This is related to sea-ice distribution excessively extending in the northern North Atlantic in MRI-CGCM3, which was greatly improved in MRI-ESM2.0 as described later. In MRI-CGCM3, warm biases were remarkable in the Southern Ocean and offshore along the western coast of continents in the Southern Hemisphere (off Peru and off Namibia). However, these errors mostly disappeared in MRI-ESM2.0. Excessive shortwave radiation in these areas in MRI-CGCM3 (Fig. 3a) and its reduction in MRI-ESM2.0 (Fig. 3b) have led to a solution for these warm SAT biases.

Warm biases in the Arctic Ocean and Canadian Archipelago are notable in MRI-ESM2.0. These biases are related to insufficient average sea-ice concentration compared to observations because of the rapid trend in decreasing Arctic sea ice since the 1980s (earlier than observations; see Section 4.3) in the historical experiment. A cold bias exceeding −2°C remains even in MRI-ESM2.0 in a small area east of Newfoundland, which is a common feature often seen in climate models that do not resolve mesoscale oceanic eddies (e.g., Keeley et al. 2012). In addition to the previously mentioned regions, there are several locations where relatively large biases remained in MRI-ESM2.0, such as cold biases over the Tibetan Plateau, Rocky Mountains, and mid-latitude North Pacific and warm biases over Central Asia and off the west coasts of South America and Africa. It is noteworthy that these are qualitatively similar to the bias distribution seen in the CMIP5 multi-model mean (e.g., Fig. 9.2b of Flato et al. 2013). Mechanisms common to many state-of-the-art climate models may contain
systematic errors.

c. **Implied ocean heat transport**

TOA radiation distribution accuracy is expected to greatly contribute to accuracy of heat transport required for the ocean by the atmosphere. Figure 6 shows implied ocean heat transport in the present-day climate, which was estimated by integrating net energy input to ocean at the surface. If the oceans are in steady state and energy is conserved in the ocean interior, this implied ocean heat transport should match with the actual heat transport by ocean processes (i.e., advection and diffusion). Compared to observation-based estimations (Fasullo and Trenberth 2008), MRI-CGCM3 underestimates northward and southward transports, particularly in the Southern Hemisphere. The southward transport exceeds 1 PW at around 10°S in the observation-based estimation, whereas in MRI-ESM2.0 it peaks at less than 0.5 PW around 20°S. This is caused by excess energy input in the mid-to-high latitudes of the Southern Hemisphere, as shown by the radiation budget bias (Fig. 3). In contrast, MRI-ESM2.0 exhibits a remarkable improvement, and meridional heat transport is perfectly consistent with observation-based estimates between 40°S and 40°N. It is also worth noting that transport across the equator is accurate.

d. **Precipitation**

Figure 7 compares the annual-mean precipitation distribution simulated by the models with observations from GPCP-1DD (Huffman et al. 2001) version 1.1. MRI-ESM2.0 captures well observed global precipitation, with a spatial correlation of 0.84. MRI-ESM2.0 shows a bias pattern in precipitation like that of MRI-CGCM3 (Figs. 7c, d). The global average precipitation excess is approximately 12% relative to the observations. Excessive precipitation in the Pacific intertropical convergence zone (ITCZ), the maritime continent, and the western equatorial Indian Ocean appear to explain a large part of this global excess.

A notable drawback in both models is the double ITCZ problem that still exists in many state-of-the-art climate models (e.g., Lin 2007). Excessively strong ITCZs, particularly in the tropical Southern Pacific are simulated along with deficient precipitation on the equator. We evaluated the severity of the double ITCZ based on the tropical precipitation asymmetry index (TPAI), which is defined as precipitation in the Northern Hemisphere tropics (equator to 20°N, area-averaged) minus precipitation in the Southern Hemisphere tropics (equator to 20°S) normalized by tropical-mean (20°S – 20°N) precipitation (Hwang and Frierson 2013). Compared to observations (TPAI = 0.207), MRI-ESM2.0 (TPAI = 0.117) shows only a moderate improvement over MRI-CGCM3 (TPAI = −0.134). Hwang and Frierson (2013) suggested that excessive energy input to the Southern Ocean is one of the causes of double ITCZ through inappropriate northward heat transport by the atmosphere. However, Kay et al. (2016) argued that reducing the Southern Ocean radiation bias has little effect on the ITCZ’s location because the change in the cross-equatorial heat transport due to heating bias reduction is dominated by the ocean rather than the atmosphere. The fact that MRI-ESM2.0’s reduction in Southern Ocean radiation bias (Fig. 3) does little to alleviate the double ITCZ supports Kay et al.’s (2016) argument.

Overall patterns of seasonal precipitation (Figs. S2, S3) are fairly well simulated in MRI-ESM2.0, though the bias patterns for each season are similar between both models. Some of the monsoon precipitation deficit in MRI-CGCM3 has been improved in MRI-ESM2.0. For example, MRI-CGCM3 had a problem with small precipitation amounts for the South Asian summer monsoon. A similar problem is also confirmed in MRI-ESM2.0, but with a slight improvement (Figs. S3c, d). Additionally, Southeast (Figs. S2c, d) and West (Figs. S3c, d) African monsoon precipitation deficits in MRI-CGCM3 are alleviated in MRI-ESM2.0. These precipitation bias differences may be due to the tuning of a few cumulus convection scheme’s parameters; however, a detailed investigation is needed.
e. **Atmospheric circulation**

Figure 8 shows meridional distributions of zonal mean zonal wind and temperature. In general, MRI-ESM2.0 reproduced more realistic zonal wind and temperature distributions than MRI-CGCM3. In the stratosphere, zonal wind biases were remarkably reduced in boreal winter (December–January–February; DJF). These biases, which were due to a weak easterly bias in the austral sub tropics and strong westerly biases due to the strong boreal polar vortex, in MRI-CGCM3 are reduced in MRI-ESM2.0 (Figs. 8a, b). In austral winter (June–July–August; JJA), strong southern polar vortex biases and northern subtropical easterly biases in MRI-CGCM3 were decreased in the upper stratosphere in MRI-ESM2.0. Figures 8e–h show zonal-mean temperature, indicating a notable reduction of warm biases in the tropical tropopause in both seasons. In addition, cold biases in MRI-CGCM3 turned to warm biases in MRI-ESM2.0 in both polar regions of the lower stratosphere (Figs. 8e, f). The increase of the easterly wind component in the summer hemispheric stratosphere is caused by the replacement of Rayleigh friction by the Hines NGWD parameterization. Furthermore, this replacement enhanced day-to-day stratospheric polar vortex variability in the winter hemisphere. The deteriorated warm bias in the whole stratosphere in MRI-ESM2.0 might be partly affected by updated solar forcing data (Matthes et al. 2017). Both enhancement of resolved wave activity due to finer vertical grid spacing in the stratosphere...
Fig. 8. Climatology (1986–2005) of zonal mean zonal wind (unit: m s$^{-1}$) for (a, b) boreal winter (December–January–February; DJF) and (c, d) boreal summer (June–July–August; JJA) and atmospheric temperature (unit: K) for (e, f) DJF and (g, h) JJA. Panels in the left and right column are for MRI-CGCM3 and MRI-ESM2.0, respectively. Color shading denotes model biases compared to JRA-55. Contour interval is 10 m s$^{-1}$ (a–d) and 10 K (e–h).
and the introduction of parameterized (NGWD) forcing accelerated stratospheric meridional overturning circulation (i.e., the Brewer–Dobson circulation) in the winter hemisphere. Intensification of the Brewer–Dobson circulation would play an important role for breaking the hemispheric-scale meridional temperature gradient bias, thus reducing tropical tropopause’s warm bias.

In the troposphere, subtropical jets of both winter hemispheres are excessively enhanced in MRI-ESM2.0 (Figs. 8b, d), whereas they were shifted equatorward in MRI-CGCM3 (Figs. 8a, c). Enhancements of the winter subtropical jets may be related to larger heating in the tropical troposphere with larger tropical-mean precipitation in MRI-ESM2.0 than in MRI-CGCM3. In both DJF and JJA, cold biases were reduced in the northern extratropical troposphere, whereas in the southern extratropical troposphere cold biases were increased. These changes reflect SAT biases, with cold biases in the Northern Hemisphere in MRI-CGCM3 being greatly reduced in MRI-ESM2.0 and SAT being relatively decreased in the Southern Hemisphere (Fig. 5).

f. Sea ice

Figure 9 shows annual cycles of simulated Northern Hemisphere and Southern Hemisphere sea-ice extent for present-day climate, together with observations of the National Snow and Ice Data Center (NSIDC) Sea Ice Index Version 3 (Fetterer et al. 2017). Both models simulate seasonal Northern Hemisphere sea-ice extent changes that are consistent with observations, in which the extent reaches a maximum in March and a minimum in September. However, MRI-CGCM3 shows a large (~25%) wintertime overestimation relative to observations, whereas the extent decreases sharply from May to August, reaching a minimum extent in September, close to the observations. In contrast, the MRI-ESM2.0 Northern Hemisphere sea-ice extent tends to be underestimated throughout the year; however, the annual cycle’s amplitude is almost the same as that of the observations (Fig. 9a). The Southern Hemisphere wintertime sea-ice extent (Fig. 9b) is slightly excessive in both models, particularly in MRI-ESM2.0. Both models simulate austral summer minimum sea-ice extents that coincide well with the observations.

MRI-ESM2.0 reasonably represents the overall geographical distribution of Northern Hemisphere sea-ice concentration in the summer and winter (Figs. 10a, b). In detail, MRI-ESM2.0 reasonably represents wintertime sea-ice edge position in the Barents Sea; however, the southward extension of the sea ice is underestimated in the Sea of Okhotsk and off the Newfoundland coast (Fig. 10a). MRI-CGCM3 wintertime sea ice overly extended in the northwestern North Atlantic and Labrador Sea (not shown, see Fig. 17 of Yukimoto et al. 2012). We speculated that this was due to insufficient mixing near the sea surface in these regions interacting with strong surface-layer stability due to erroneously low salinity compared with observations.

Although the sea-ice model component itself was not changed between MRI-CGCM3 and MRI-ESM2.0, the vertical surface-layer mixing was strengthened in the MRI-ESM2.0 ocean model compared to the MRI-CGCM3 ocean model because of vertical mixing scheme changes (Section 2.4). This has led to a realistic stratification near the sea surface and a sea-ice
distribution closer to observed values. The summertime sea-ice extent in the entire Arctic Ocean is lower than that in the observations for 1986–2005, and the climatology is closer to the distribution observed for the more recent decade (2006–2015; Fig. 10b). It should be noted that Northern Hemisphere sea ice is currently in a process of rapid decrease and is thus sensitive to choices of averaging period. In MRI-ESM2.0, the SAT shows a warm bias in the Arctic region (Fig. 5b) mainly due to a marked wintertime warm bias (Fig. S1). This is consistent with underestimation of Arctic sea ice resulting in excessive heat release from the open sea surface or through the thin sea ice. The sea-ice edge distribution of the Antarctic sea-ice distribution simulated by MRI-ESM2.0 (Figs. 10c, d) is reproduced fairly well.

g. Ocean surface

The annual-mean SST bias in MRI-ESM2.0 relative to COBE-SST2 observations (Hirahara et al. 2014) is compared to that in MRI-CGCM3 (Figs. 11a, b). Large warm biases were observed in MRI-CGCM3
over the Southern Ocean and eastern South Pacific, but such biases were not found in MRI-ESM2.0. Those SST biases possibly result from improvements in simulating downward shortwave radiation (Fig. 3). Cold biases in the Kuroshio Extension region are significantly lower in MRI-ESM2.0 than in MRI-CGCM3. This improvement might be related to the reduced sea-surface height (SSH) bias discussed later in this section. MRI-CGCM3 also exhibited large cold biases in the Labrador Sea and the Greenland-Iceland-Norwegian (GIN) seas, which resulted from excessive wintertime sea-ice extent, as mentioned above. MRI-ESM2.0 reproduces well the wintertime Northern Hemisphere sea-ice distribution (Fig.10), probably due to the upgraded mixed layer scheme, as mentioned previously. This accompanies a significant SST
bias reduction in MRI-ESM2.0. MRI-ESM2.0 SST values show small biases below 1°C around the equator in the Pacific, whereas slightly cold biases appear in MRI-CGCM3 because of the overestimated cold tongue.

The simulated annual mean SSS is compared to version 2 dataset of the World Ocean Atlas 2013 (WOA13v2; Zweng et al. 2013; Figs. 11c, d). The observed SSS is calculated as an average of two periods, 1985–1994 and 1995–2004, and simulated SSS is an average for 1985–2004. MRI-CGCM3 shows prominent fresh biases around the North Atlantic Current, which suppresses deep convection in the North Atlantic. Fresh biases exist east of Newfoundland in both models, though the magnitude and extent are largely reduced in MRI-ESM2.0. In fact, MRI-ESM2.0 reproduces well Atlantic thermohaline circulation, as shown later. Other strong fresh biases around the Antarctica and in the subtropical South Pacific in MRI-CGCM3 are notably reduced in MRI-ESM2.0. The subtropical South Pacific bias reduction could be possibly accounted for by MRI-ESM2.0 alleviating excessive evaporation from MRI-CGCM3 (Fig. 7). A remarkable saline bias is observed in the Bay of Bengal in both models due to insufficient precipitation, but its magnitude is slightly decreased in MRI-ESM2.0 because of improved Asian summer monsoon precipitation (Fig. 7) and associated river runoff. In the Arctic, significant biases on the Atlantic side are not seen in MRI-ESM2.0, unlike in MRI-CGCM3; however, Pacific side saline biases worsened in MRI-ESM2.0. For causes of the Arctic saline biases, we examined several candidates, such as relationship with the river runoff, but we could not identify any clear cause. Detailed budget analysis may be required to identify the causes.

The simulated annual-mean SSH from 1993 to 2005 is compared to the annual-mean absolute dynamic topography provided by the Copernicus Marine Environment Monitoring Service (CMEMS; Figs. 11e, f). Simulated SSHs averaged globally from 60°S to 60°N are adjusted to those of CMEMS absolute dynamic topography on a climatological basis. MRI-CGCM3 exhibits large negative biases in the North Pacific subtropical gyre, which implies a weak western boundary current that likely yields to less heat transport to the Kuroshio Extension region, causing a cold SST bias there (Fig. 11a). Similar negative biases are seen in MRI-ESM2.0 but are much smaller because of the stronger subtropical gyre due to amplified zonal wind stress associated with westerlies and trade winds over the North Pacific (not shown). These are consistent with accompanying SST biases. Large positive SSH biases in dense water formation regions, such as the Labrador, Ross, and Weddell seas, are observed in MRI-CGCM3 but are not significant in MRI-ESM2.0. These improvements are likely related to better deep and bottom water formation process representations as discussed below. Large SSH biases are found in both models around recirculation gyres of the Gulf Stream, the Kuroshio Current, and the Brazil Current extension regions, which are common in many CMIP5 models (Landerer et al. 2014) employing ocean models with low horizontal resolution.

h. Ocean interior

Figure 12 presents depth-latitude sections of zonal and annual average potential temperature and salinity climatology and their biases from WOA13v2 climatology (Locarnini et al. 2013; Zweng et al. 2013). Southern Ocean bottom waters in MRI-CGCM3 have remarkable warm biases exceeding 1°C, as discussed in Yukimoto et al. (2012). In MRI-ESM2.0, similar but slightly warm biases appear, whereas weak fresh biases in salinity appear (Fig. 12d). Both models failed to properly represent bottom water formation processes around Antarctica. Large warm biases north of 60°N are larger in MRI-ESM2.0 than MRI-CGCM3, which can be accounted for mostly by warm biases beneath the subsurface in the GIN seas (not shown), similarly to MRI-CGCM3 (Yukimoto et al. 2012). Additionally, Atlantic thermohaline circulation is remarkably stronger in MRI-ESM2.0 than MRI-CGCM3 and accompanies a larger northward heat transport in the North Atlantic (Fig. 6), which probably resulted in warming in the GIN seas.

MRI-ESM2.0 and MRI-CGCM3 both reproduce overall structures of meridional overturning circulation (MOC; Fig. 13) well, but there are quantitative differences between the two models. In MRI-ESM2.0, the upper limb of the Atlantic MOC associated with North Atlantic Deep Water formation is strongly enhanced. The temporal average for 2004–2008 amounts to 17.7 Sv, which is comparable with the observation-based estimate of 18.7 Sv from RAPID-MOC (Rayner et al. 2011). This Atlantic MOC enhancement may contribute to the more realistic northward heat transport in the Northern Hemisphere (Fig. 6).

The subpolar cell around Antarctica was far offshore (around 64°S) in MRI-CGCM3. The winter mixed layer is shallow in the coastal areas around Antarctica and unrealistically deep at low latitudes in the offshore region of the East Antarctica (Figs. S4a, b). These indicated that dense water formation did not take
place in coastal polynyas, and relatively warm bottom waters were created by open ocean convection in this low-latitude region that resulted in warm biases in the bottom layer of the Southern Ocean in MRI-CGCM3. In contrast, MRI-ESM2.0’s subpolar cell extends to higher latitudes (its core locates at 70°S) and the contour line of zero reaches the sea surface. Deep mixed layers in MRI-ESM2.0 are limited to coastal areas around Antarctica in austral winter (Fig. S4c). These imply that dense water realistically formed in coastal polynyas around Antarctica drives the subpolar cell in this model. This improved representation of bottom water formation resulted in reduced SSH bias around Antarctica (Figs. 11e, f) and cold bottom waters that relieved warm biases in the bottom layer of the Southern Ocean (Figs. 12a, c). Meanwhile, subpolar cell strength was found to be weak in MRI-ESM2.0, which indicated insufficient dense water formation around

Fig. 12. Biases (color shading) of zonal and annual-mean potential temperature (left) and salinity (right) relative to WOA13v2 1955–2012 mean in (a, b) MRI-CGCM3 and (c, d) MRI-ESM2.0. Contour lines denote the model climatologies. Averages of model outputs from 1986 to 2005 are shown.
Antarctica. It is consistent with the warm and fresh bias of Southern Ocean bottom water (Figs. 12c, d). In addition, the bottom limb of the MOC in the Southern Ocean and the Indo-Pacific MOC associated with the Circumpolar Deep Water (CDW) formation are extensively weakened in MRI-ESM2.0. Roemmich et al. (1996) estimated northward CDW transport through the Samoan Passage and surrounding areas between Fiji and Tahiti at 10.6 ± 1.7 Sv. MRI-CGCM3 shows a comparable magnitude (12.1 Sv) for maximal northward CDW transport at 30°S, but in MRI-ESM2.0 it only amounts to 5.6 Sv. This large decrease might have partially resulted from changes to the background vertical diffusion coefficient and insufficient Southern Ocean bottom water formation.

1. El Niño Southern Oscillation
The El Niño Southern Oscillation (ENSO) is major climate variability in the tropical Pacific, which has a seasonal to interannual time scale and significantly

Fig. 13. Residual meridional overturning circulation (MOC) for the Southern Ocean (left), Atlantic Ocean (middle), and Indo-Pacific Ocean (right) in (a) MRI-CGCM3 and (b) MRI-ESM2.0.
affects regional climates around the world through various teleconnections. There is a great interest in how ENSO and its consequences will change due to future climate change (e.g., Christensen et al. 2013); therefore, it is important that climate models used to project climate change can reproduce realistic ENSO variability in space and time.

Figures 14a–c show SST anomaly distributions for the observations and models based on a regression on the Niño-3 (the region of 150°W–90°W, 5°S–5°N) SST anomaly time series. MRI-CGCM3’s SST anomaly distribution (Fig. 14b) resembles the observations (Fig. 14a) except for its magnitude, which is generally smaller. The MRI-ESM2.0 SST anomaly regression map (Fig. 14c) is closer to the observations than that of MRI-CGCM3; however, its magnitudes are slightly larger than the observations.

Not only has the SST anomaly become more realistic in MRI-ESM2.0, but the precipitation and teleconnection responses have improved (Figs. 14d–f). Precipitation responses to ENSO (e.g., increased central equatorial Pacific precipitation and a decrease in surrounding regions including the Western Pacific during El Niño) are realistically reproduced by MRI-ESM2.0 (though slightly overestimated in some areas), whereas these signals were weak in MRI-CGCM3. MRI-ESM2.0 also represents remote precipitation anomalies of a decrease over equatorial South America and increases over tropical eastern Africa, southeastern China through southern Japan, and the United States. In MRI-ESM2.0, the mean sea-level pressure response is also realistic, with a zonal seesaw pattern in the tropical Pacific accompanied by a remote signal in the northeastern North Pacific that is very similar to the observed pattern. This implies that ENSO-related atmospheric teleconnections are also realistic in MRI-ESM2.0. We consider that an improved central equatorial Pacific precipitation response led to realistic atmospheric teleconnections.

MRI-CGCM3 simulated a smaller ENSO amplitude than observed values. Standard deviations of the Niño-3 SST anomaly were compared between models (in the historical experiments) and observations from 1979 to 2005. The standard deviation of 0.60°C in MRI-CGCM3 is significantly smaller than that of 0.96°C for the observations (COBE-SST2; Hirahara et al. 2014). In MRI-ESM2.0, the standard deviation of the Niño-3 SST is 1.04°C, which is sufficiently close to the observations. In addition, compared to MRI-CGCM3, MRI-ESM2.0 has a Niño-3 SST power
spectrum (Fig. 15) closer to that of the observations, which show wide-band power at periods from 2 to 7 years, though it is slightly overestimated between 2 and 4 years. MRI-CGCM3 has a similar structure of Niño-3 SST temporal variations to MRI-ESM2.0, but has much smaller power spectra in the whole range for periods shorter than 50 years. The reason why ENSO got stronger in MRI-ESM2.0 is complicated. Presumably, many changes are involved, including changes to the mixed layer scheme and some ocean model parameters as well as those of the cloud scheme and some parameters of cumulus convection of the atmosphere model. Further investigation is needed to determine the cause.

j. Stratospheric QBO

The vertical resolution enhancement in the upper troposphere and stratosphere, which enables weakened background vertical diffusion, and implementing NGWD parameterization contributed crucially to drive the realistic QBO. Figure 16 shows a comparison of equatorial zonal mean zonal wind variations for JRA-55 and MRI-ESM2.0. MRI-ESM2.0 simulates realistic QBO structures, with downward propagation of alternating easterly and westerly zonal winds in the equatorial stratosphere. The bottom edges of the

Fig. 15. Power spectra of the Niño-3 sea-surface temperature (SST) anomaly for observations (COBE-SST2; black), MRI-CGCM3 (blue), and MRI-ESM2.0 (red) for the period 1850–2005. Thin dashed lines indicate individual members of the ensemble simulations and thick lines indicate ensemble means.

Fig. 16. Time-pressure cross section of the zonal mean zonal wind averaged over 5°S–5°N in (a) JRA-55 and (b) MRI-ESM2.0. Contour interval is 8 m s\(^{-1}\).
downward propagation reach well into the lower stratosphere. The simulated QBO period is 30 months, close to the 28-month period of JRA-55, and the simulated wind variation amplitude is also realistic, but slightly larger at around 30 hPa than that of the reanalysis. MRI-CGCM3 did not simulate any QBO-like variation in the stratosphere (not shown).

4.3 Historical climate change
A climate model faithfully reproducing past climate change is one of the important factors measuring the reliability of future climate change projections. To this end, a model is required to present appropriate radiative forcing and climate sensitivity. Figure 17 compares the global-mean SAT of the MRI-CGCM3 and MRI-ESM2.0 historical experiments with observations (HadCRUT4; Morice et al. 2012). Model experiment plots are aligned so the 1851–1900 averages agree with the observations to more distinctly display SAT changes relative to the preindustrial level. Because the plots are 5-year running filtered five-member ensemble averages, interannual fluctuations as internal variability (except for responses to major volcanic eruptions such as Mt. Agung in 1963 and Mt. Pinatubo in 1991) are smaller than the observations. Both models reasonably reproduce increases in SAT on a secular time scale, and qualitatively reproduce the observed warming trend in SAT during the first half of the 20th century and the subsequent weak cooling trend. However, in the MRI-ESM2.0 experiment, SAT began to decrease at a greater rate than that observed for the 1950s through the 1960s. Such rapid SAT cooling is not seen in the MRI-CGCM3 experiment, in which an underestimated temperature change is displayed throughout the 20th century. Both models reproduce the relatively rapid temperature increase from the 1970s onward; however, there are differences in trend magnitude. The trend for the 30-year period of 1976–2005 is underestimated (1.14°C/century) in MRI-CGCM3 and overestimated (2.16°C/century) in MRI-ESM2.0 compared to the observed trend (1.96°C/century). As a result, the MRI-ESM2.0 experiment has a temperature deviation at the beginning of the 21st century relative to the preindustrial level that is close to observations, whereas the deviation is underestimated in the MRI-CGCM3 experiment.

From the 1950s to the 1960s, the MRI-ESM2.0 experiments showed a marked increase in the simulated sulfate aerosol optical depth (Fig. S5), implying that the negative radiative forcing of sulfate aerosol become stronger than before. Studies comparing and evaluating effective radiative forcing in models separately for aerosols and greenhouse gases are ongoing to participate in the Radiative Forcing Model Intercomparison Project (RFMIP; Pincus et al. 2016), and it is anticipated that the results will address the question of whether forcings are properly prepared for reproduction of past climate. The difference in the warming trend since the 1970s between the MRI-CGCM3 and MRI-ESM2.0 experiments suggests a difference in climate sensitivity to increased CO2 between the models, as the positive radiative forcing of CO2 and other greenhouse gases predominates over negative radiative forcing from aerosols during this period. Climate sensitivity has increased from 2.6 K in MRI-CGCM3 to 3.1 K in MRI-ESM2.0, which was estimated from separate experiments with abruptly quadrupled CO2. Detailed investigations on climate sensitivity differences between the models are also in preparation.

The summer sea-ice extent is significantly affected under global warming conditions. Figure 18 shows historical changes in the summer sea-ice extent in the Northern and Southern hemispheres. The Northern Hemisphere sea-ice extent reproduced by MRI-ESM2.0 shows a rapid decrease since the mid-1980s.
from a preindustrial era level of ~ 7 million km\(^2\). The simulated rapid sea-ice decrease seems to be about 10 years earlier than the observed decrease that started in the mid-1990s. This led to an underestimation of the sea-ice extent (Fig. 9a) in MRI-ESM2.0 under present-day climate conditions (average of 1986–2005). The MRI-CGCM3 sea-ice extent time series shows a gradual declining trend over the entire 20th century, with interdecadal variation after the 1960s, which seems different from the observed behavior during the satellite era. As Stroeve et al. (2012) suggested models showing a rapid decrease in summer sea-ice extent tend to have relatively thinner sea-ice thickness at the beginning of the historical experiment. In the piControl simulation of MRI-ESM2.0, the maximum sea-ice thickness near the North Pole was about 2.5 m, whereas it was more than 3 m in MRI-CGCM3.

The Southern Hemisphere summer sea-ice extent simulated by MRI-CGCM3 was almost constant until it started decreasing in the 1990s. In the MRI-ESM2.0 experiment, the sea ice gradually decreased throughout the 20th century. Neither model reproduces the weakly increasing observed trend. Most CMIP5 models also do not reproduce the increasing trend of sea-ice area seen in the satellite observations (Zunz et al. 2013). The data period is likely too short to determine whether this increasing trend is within the range of natural variability or response to external forcing (Jones et al. 2016b). It is also pointed out that the strengthening of the Southern annular mode associated with the ozone depletion may strengthen the westerly wind which may increase sea ice. However, it is also uncertain whether the models can properly express the mechanism (Jones et al. 2016b).

Figure 19 shows a time series of the Atlantic meridional overturning streamfunction at 26.5°N and a depth of 950 m for MRI-CGCM3 and MRI-ESM2.0. MRI-ESM2.0 shows a 2.5 to 7.3 Sv increase in magnitude over MRI-CGCM3. The Atlantic MOC in MRI-ESM2.0 shows a local maximum around 1990. Such a local maximum is also found in a preliminary analysis of the Detection and Attribution Model Intercomparison Project (DAMIP; Gillett et al. 2016) experiments that include a separate historical experiment with only aerosol forcing in MRI-ESM2.0. The local maximum in the historical experiments in MRI-ESM2.0 might be associated with the enhanced anthropogenic aerosol forcing in this period (Fig. S5) and its resultant cooling over the North Atlantic. Detailed investigation is needed in a future study.

5. Summary and conclusions

The new Earth system model MRI-ESM2.0 has been developed at MRI. This model is based on the previous models, MRI-CGCM3 and MRI-ESM1, that participated in CMIP5, with numerous improvements meant to provide highly accurate climate simulations. We described updates of model formulations and evaluated the basic performance of its physical component. The model is an integration of interactive models for aerosol and atmospheric chemistry into the atmosphere–ocean coupled model. The number of
vertical layers in the atmospheric models has been increased from 48 to 80, and the vertical resolution has been enhanced in the tropopause through the stratosphere–mesosphere and near the surface. In addition, a NGWD parameterization and improved stratocumulus cloud scheme have been introduced. Numerous improvements have been added to processes, such as cloud macro- and micro-physics, cloud radiation, aerosols, and atmospheric chemistry. The ocean–sea-ice model has also been updated, implementing a generic length scale scheme for vertical mixing. In contrast, there were no changes in atmospheric horizontal resolution and components such as atmospheric radiation code, atmospheric turbulence scheme, land surface model, and terrestrial and ocean carbon cycles.

Following a preindustrial spin-up, we conducted a preindustrial control experiment and a set of historical simulations from the mid-19th century to the present, which were driven by forcing based on observations. These experiments all conformed to the protocol specified by CMIP6. After the 1000-year spin-up, climate drift is small, at least near the surface, with an acceptable value less than 0.02°C/century for global-mean SAT and SST.

The reproducibility of present-day climate in the historical experiment was evaluated and compared with observations and reanalysis for the same period. For SAT, Northern Hemisphere cold bias and Southern Hemisphere warm bias, which were noticeable in MRI-CGCM3, were remarkably reduced in MRI-ESM2.0 and showed good fidelity. Improvements in cloud representation significantly contributed to a reduced SAT bias (particularly the warm bias in the Southern Ocean). The accumulation of various improvements concerning clouds, including the new stratocumulus cloud scheme, led to remarkably reduced errors in shortwave, longwave, and net radiations at TOA, becoming sufficiently small compared with those in the CMIP5 models. Improved radiation distribution means that the meridional heat transport required for the ocean coincides well with observation-based estimates.

MRI-ESM2.0 displays realistic reproduction in both mean climate and interannual variability. For instance, stratospheric QBO can now be realistically expressed because of the enhanced vertical resolution and introduction of NGWD parameterization. In addition, simulated ENSO in MRI-ESM2.0 has become closer to the observed ENSO than that in the previous model. MRI-ESM2.0 reasonably reproduces global SAT changes in recent decades, though the interdecadal change in the trend is overestimated compared with observations. Although MRI-ESM2.0 reproduces the observed rapid decrease of Arctic sea-ice extent in recent decades, the decrease started too early, resulting in an underestimation of present-day sea-ice extent.

MRI-ESM2.0 has been improved over the previous models MRI-CGCM3 and MRI-ESM1 in many aspects, and the new model is expected to demonstrate superior performance in the CMIP6 experiments. However, it still has several drawbacks, such as a double ITCZ bias in the tropical precipitation distribution. With MRI-ESM2.0, we are participating in various satellite model intercomparison projects (MIPs) of CMIP6, including the Scenario MIP (ScenarioMIP; O’Neill et al. 2016) and the Decadal Climate Prediction Project (Boer et al. 2016) for projecting or predicting the future climate, and the Cloud Feedback MIP (CFMIP; Webb et al. 2017), DAMIP (Gillett et al. 2016), RFMIP (Pincus et al. 2016), the Ocean MIP (OMIP; Griffies et al. 2016), the Aerosol Chemistry MIP (AerChemMIP; Collins et al. 2017), the Flux-Anomaly-Forced MIP (FAFMIP; Gregory et al. 2016), the Global Monsoons MIP (GMMIP; Zhou et al. 2016),
2016), and the Coupled Climate-Carbon Cycle MIP (C4MIP; Jones et al. 2016a) for exploring response to the forcing and feedbacks and for investigating causes of systematic biases. Through vast experiments and analysis of these projects, we can expect advancement of understanding of the climate system and improvements of the model.

Supplements

Supplement 1 shows distributions of seasonal-mean SAT bias (relative to JRA-55 reanalysis, 1986–2005 mean) for the present-day DJF and JJA climatologies in MRI-CGCM3 and MRI-ESM2.0. Supplements 2 and 3 show same as Fig. 7 except for DJF and JJA means. Supplement 4 shows observed and simulated mixed layer depth in the Southern Ocean in September. Supplement 5 shows temporal evolutions of the globally averaged annual-mean aerosol optical depth in the historical experiments for each aerosol species in MRI-ESM2.0 and total aerosol optical depth in MRI-CGCM3.

Acknowledgments

This work was undertaken as part of the Research Program on Climate and Environmental Change with Advanced Climate Modeling funded by MRI, Japan Meteorological Agency. The computation for this work was performed on a Fujitsu FX100 supercomputer at MRI. This work was partly supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI (grant numbers: JP26701004, JP15H05816, JP16H01772, JP18H03363, and JP18H05292), the Environment Research and Technology Development Fund (2-1703 and S-12) of the Environmental Restoration and Conservation Agency, Japan, the Integrated Research Program for Advancing Climate Models (C4MIP; Jones et al. 2016a) for exploring response to the forcing and feedbacks and for investigating causes of systematic biases. Through vast experiments and analysis of these projects, we can expect advancement of understanding of the climate system and improvements of the model.

Supplements

Supplement 1 shows distributions of seasonal-mean SAT bias (relative to JRA-55 reanalysis, 1986–2005 mean) for the present-day DJF and JJA climatologies in MRI-CGCM3 and MRI-ESM2.0. Supplements 2 and 3 show same as Fig. 7 except for DJF and JJA means. Supplement 4 shows observed and simulated mixed layer depth in the Southern Ocean in September. Supplement 5 shows temporal evolutions of the globally averaged annual-mean aerosol optical depth in the historical experiments for each aerosol species in MRI-ESM2.0 and total aerosol optical depth in MRI-CGCM3.

Acknowledgments

This work was undertaken as part of the Research Program on Climate and Environmental Change with Advanced Climate Modeling funded by MRI, Japan Meteorological Agency. The computation for this work was performed on a Fujitsu FX100 supercomputer at MRI. This work was partly supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI (grant numbers: JP26701004, JP15H05816, JP16H01772, JP18H03363, and JP18H05292), the Environment Research and Technology Development Fund (2-1703 and S-12) of the Environmental Restoration and Conservation Agency, Japan, the Integrated Research Program for Advancing Climate Models (C4MIP; Jones et al. 2016a) for exploring response to the forcing and feedbacks and for investigating causes of systematic biases. Through vast experiments and analysis of these projects, we can expect advancement of understanding of the climate system and improvements of the model.


Kawatani, Y., K. Sato, T. J. Dunkerton, S. Watanabe, S. Miyahara, and M. Takahashi, 2010: The roles of equa-


Nagasawa, R., 2012: The problem of cloud overlap in the radiation process of JMA’s global NWP model. CAS/JSC WGNE Research Activities in Atmospheric and Oceanic Modelling/WM0, 4, 4.15–4.16.


Smolarkiewicz, P. K., 1984: A fully multidimensional positive definite advection transport algorithm with small


