

Response to anonymous referee #2

Major Comments:

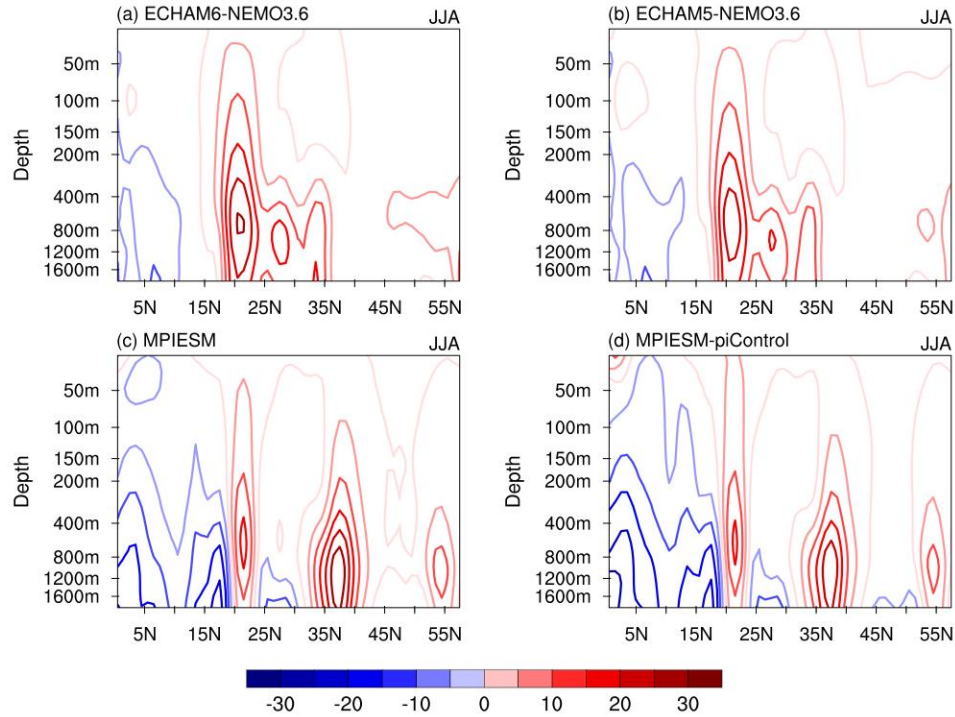
*1. The authors compared mostly the Pacific Ocean and concentrated on tropical Pacific. One of the stark improvements lie in the North Atlantic, especially where the cold bias in the subpolar region seen in MPI-ESM has been ameliorated. It would be extremely beneficial for the community to discuss this too, and show how the AMOC may be different between these different CGCMs.*

Reply:

Thank you for these comments. We have compared the AMOC simulation results of the three CGCMs. The inter-model differences imply that cold SST in North Atlantic is partially attributed to weak AMOC, which is consistent with previous studies. The experiment results have been introduced in section 5.1. The MOC bias figures have been redrawn with the same color settings that are easier to recognize.

‘Since the upper cell of Atlantic meridional overturning circulation (AMOC) plays a significant role in delaying warming signals from anthropogenic greenhouse gases and responding to climate change (Marshall et al., 2014; Buckley and Marshall, 2016), model bias analysis is still focused on the upper ocean levels. The overall magnitude of AMOC bias is less than that of NPMOC with significantly reduced biases near the sea surface (Fig. 10), which is consistent with those of surface currents among the three CGCMs. The ECHAM6-NEMO3.6 shows exiguous bias near the ocean surface, but presents strong biases in the mesopelagic zone of subtropical areas, bringing more heat to higher latitudes (Fig. 10a). Likewise, the ECHAM5-NEMO3.6 exhibits strong circulation biases rotating clockwise in the thermocline that intensifies poleward heat transport (Fig. 10b). With similar bias patterns of the AMOC, the ECHAM5-NEMO3.6 and the ECHAM6-NEMO3.6 have opposite SST biases in North Atlantic (Figs. 3a and 3c), which implies that the air-sea feedback including WES feedback and NAM as suggested by Zhang & Zhao (2015) takes the responsibility. The MPI-ESM experiment shows negative biases in tropical Atlantic from the sea surface to the bathypelagic zone, indicating that the overturning circulation has been restrained. There is a narrow positive bias in the subtropical Atlantic, but its strength has been limited by the negative biases nearby. One consequence of the weak AMOC is the decrease of SST in North Atlantic due to less heat supply from the tropics (Fig. 3e). The overturning circulation is enhanced in the middle latitudes with one centre located north of 35°N and another centre around 55°N at the depth of 1200m. It still promotes the poleward heat transport and results in warm SST biases in subpolar region (Fig. 3e). The AMOC biases in the MPI-ESM piControl experiment are similar as those in the MPI-ESM experiment, with more negative biases in tropical Atlantic. Comparing the AMOC biases between the MPI-ESM and the ECHAM5-NEMO3.6, it can be seen that the SST cold biases in North Atlantic are partially attributed to decreased MOC in the thermocline of tropical and extra-tropical oceans. However, the air-sea interaction also takes account of the SST variations in consideration of the SST

differences between the ECHAM5-NEMO3.6 and the ECHAM6-NEMO3.6. Zhang & Zhao (2015) suggested that the cold SST bias in Atlantic caused the same cold bias in North Pacific through different mechanisms originating in tropical and extra-tropical Atlantic. Because the differences of NPMOC are bigger than those of AMOC between these two newly developed CGCMs, it suggests an inverse cause-and-effect relationship between the cold SST biases in North Pacific and North Atlantic where the former takes the lead.



**Figure 10: Model biases of AMOC in summer, (a) ECHAM6-NEMO3.6, (b) ECHAM5-NEMO3.6, (c) MPI-ESM, (d) MPI-ESM piControl, Unit: Sv.'**

*2. Changes in mean state from changing component model provide simple but necessary information. However, it would benefit readers and the community if more information such as variability of the system is assessed, such as variance of variables that the authors have covered, and maybe even power spectrum of ENSO and SAM.*

Reply:

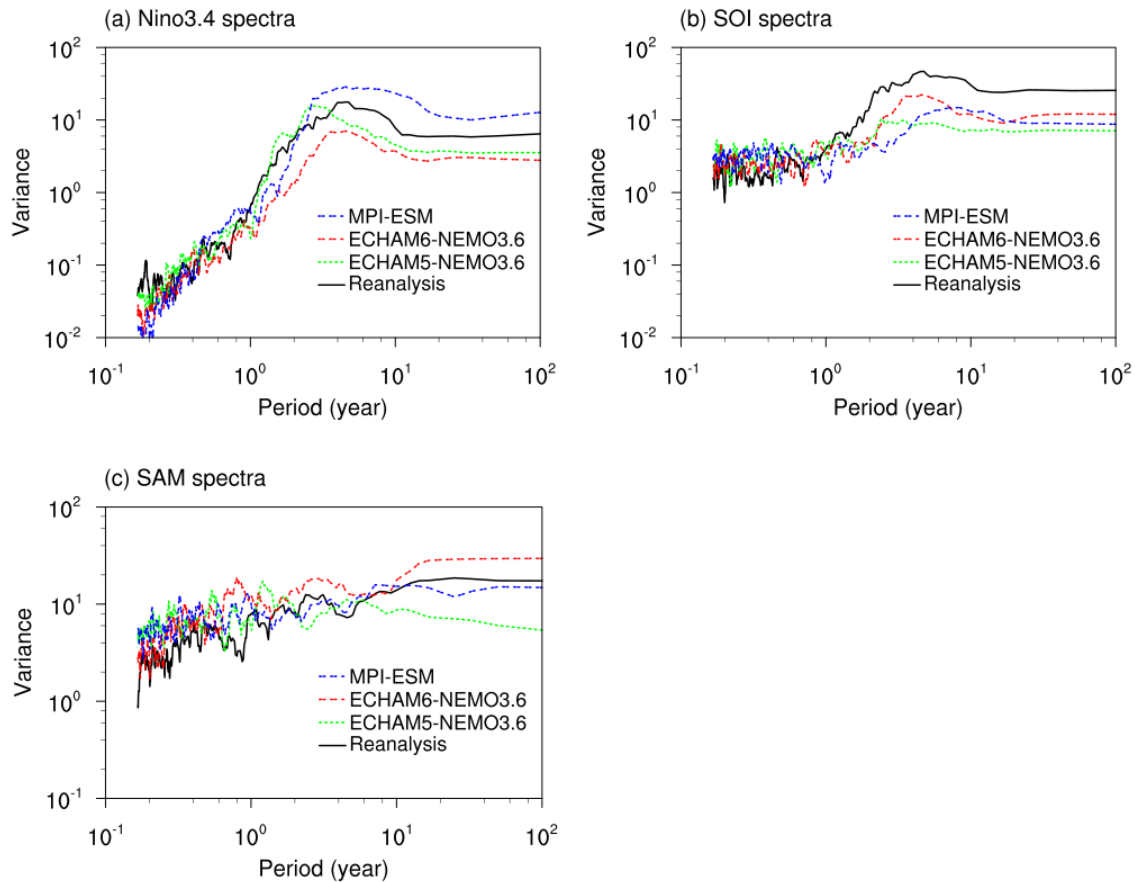
Thanks for your comments. We have added assessment of the system variability, including power spectrum analysis of Niño3.4 index, SOI and SAM index. Because the model data in use spans 100-year realizations, much longer than that of the ERA-Interim reanalysis, the ERA-20c has been used instead to evaluate the model variability. The SST reanalysis data from Hadley center (HadISST) is still used for SST variability assessment.

Both HadISST and ERA-20c reanalyses have 100-year time span from 1911 to 2010. Model variability evaluation has been provided in section 4.7.

#### **‘4.7 Model variability of ENSO and SAM**

In the coupled ocean-atmosphere system, global climate variability has been driven by the El Niño-Southern Oscillation (ENSO), the southern annular mode (SAM, also called the Antarctic Oscillation) and the Indian Ocean dipole (IOD) (Philander, 1990; Wallace and Thompson, 2002; Saji et al., 1999). It is therefore necessary to examine the model variability by applying spectra analysis on relevant indices. The CGCM simulations of the three indices are generally consistent with the theoretical red noise (Markov) spectrum (figure omitted). The Niño3.4 index is defined as the SST anomalies averaged over the NINO34 region (5°N-5°S, 170°W -120°W). It shows high variance in 2-7 years' period that documents the ENSO peaks in the HadISST reanalysis (Fig. 8a). All of the CGCMs reproduce similar variations of the Niño3.4 power spectra. The ECHAM6-NEMO3.6 presents weak variabilities at the interannual and interdecadal scale, whose periodic peaks are about one year less than the reanalysis counterpart. The ECHAM5-NEMO3.6 shows a better spectral distribution that best coincides with the reanalysis at the interannual scale. However, it still suffers a weak variability at the interdecadal scale and the periodic peak is even half a year less than that of the ECHAM6-NEMO3.6. The MPI-ESM instead takes on an intensified interannual variability, which stays strong at the interdecadal scale. The Southern Oscillation Index (SOI) is calculated based on the differences in sea level pressure anomalies between Tahiti and Darwin in Australia. In comparison to the Niño3.4 spectra, the SOI exhibits similar peaks at the interannual scale in the ERA-20c reanalysis (Fig. 8b). Nevertheless, all the CGCMs reproduce weak variabilities at the interannual and interdecadal scales. The ECHAM6-NEMO3.6 presents the best simulation with a significant increase in variance around 4 years' period, while the ECHAM5-NEMO3.6 shows the weakest variability at the interannual scale. It implies that the AGCM replacement has an opposite effect on the Niño3.4 and SOI variabilities. The MPI-ESM also show reduced variance from the annual scale and above, quite the opposite to that in the Niño3.4 case. Since the model biases of ENSO variability may be attributed to thermocline feedback and zonal wind variations (Borlace et al., 2013), the reversed changes in the variabilities of Niño3.4 and SOI can be caused by the related oceanic and atmospheric processes. The SAM index is calculated following Gong and Wang (1999) by the differences of normalized monthly zonal mean sea level pressure at 40°S and 65°S. Variations of SAM tend to be more flattened than those of SOI in the ERA-20c reanalysis (Fig. 8c), with prominent fluctuations from biannual to interannual scales. Compared with the reanalysis counterpart, all CGCMs show more power at interannual time scales that represents a robust modulation of the SAM, which is possibly attributed to the semi-annual oscillation (SAO) (Hurrell and van Loon, 1994) and circulation anomalies over Antarctica (Thompson and Solomon, 2002). The ECHAM6-NEMO3.6 presents stronger decadal variability than that of the reanalysis data, while the ECHAM5-NEMO3.6 exhibits weaker low-frequency variability. Since the high variance in low-frequency band represents the upward trend of SAM index at decadal scale (Raphael and

Holland, 2006), updating the AGCM can result in a drastic change of long-term climate variability in southern hemisphere. In contrast, the MPI-ESM shows the SAM variability very close to the reanalysis counterpart from 1 year and above, indicating that the OGCM feedback to the atmosphere can lead to a better representation of the inter-decadal variability.



**Figure 8:** Power spectra of (a) Niño3.4 index, (b) SOI, (c) SAM index. Solid line denotes the calculation results of reanalysis data, green dotted line denotes the ECHAM5-NEMO3.6 simulation, red dotted line denotes the ECHAM6-NEMO3.6 simulation, and blue dotted line denotes the MPI-ESM simulation.'

3. Section 6.2: Please be aware that correlation does not indicate intensity or magnitude of relationship. Fig. 13 shows that shortwave radiation has a much large magnitude and can thus possibly have a greater effect on SST even though it has a smaller correlation value compared to sensible heat and longwave radiation. This does not change the proposed mechanism, but SW radiation should be incorporated into the explanation.

Reply:

Thank you for these comments. We have rewritten this section to explain the model differences with a full picture of the air-sea interactions. Relevant content in section 7 has also been revised. Major modifications are pasted below.

‘... Since the top 3 variables that are most relevant to SST deviations (Tab. 1) suggest changes in the surface heat budget, involving the momentum and temperature exchanges between ocean and atmosphere, it implies that the joint effects of atmosphere and ocean models lead to the model deviation patterns (Fig. 15). Simulation of ocean dynamics is different after replacing the OGCM. With changes in ocean advection, the SST and surface currents are altered which modulates the surface evaporation, convection and heat conduction. Subsequently the latent heat and sensible heat fluxes vary over the sea surface. The thermal and moisture perturbations from the ocean are passed to the atmosphere during coupling processes (Fig. 15c, d). Variations of low-level atmospheric circulation and humidity take effects on cloud formation and cloud liquid water path that changes precipitation and cloud radiative forcing. The net shortwave and longwave radiations are influenced and make a difference to the atmospheric circulation (Fig. 12c) and the heat budget over the sea surface (Fig. 15 a, b). Then, the perturbation signal is transferred back to the ocean that changes the SST and surface currents. This air-sea feedback finally reaches a quasi-equilibrium with marked SST warming over vast maritime spaces over the globe. The associated physical processes represented by each radiation term in Figure 15 are in accordance with their signs of pattern correlation (Tab. 1). It seems to suggest that surface evaporation plays a predominant role in the SST differences, because the latent heat flux holds the biggest correlation coefficient and its deviation pattern is the most obvious among the four radiation terms. However, it is more likely a manifestation of the large-scale ocean dynamical effect on the inter-model differences as suggest by Ying and Huang (2016).

... So far, the analysis has made it clear that changing the OGCM affects the SST simulation through large-scale air-sea feedback that mainly involves surface evaporation, heat conduction, atmospheric convection and cloud masking of incoming and outgoing radiative fluxes. Under the impact of net radiation surplus over western Pacific, the temperature rises with upward motion which forms a warm advection heading east and descending over eastern Pacific....’

*4. Along the lines of pattern correlation being used to rank the importance of contributing factors. Since the region of focus is the Pacific, the pattern correlation should does be performed over the Pacific rather than globally. This should not change the concluding results, but it may provide some differences in the correlation and add robustness to your estimates.*

Reply:

Thank you for these comments. We have added the correlation analysis over the Pacific to facilitate the discussion of related physical processes in section 6. As you suggested before, the assessment of SAM index variability and AMOC have been added to the paper. The focus of the paper has thus been expanded to other oceans, and we still keep the previous correlation analysis performed at the global scale in the manuscript. Major modifications are as follows:

‘... Because the analysis in previous sections is mainly focused on the Pacific, the pattern correlation over the area is also provided for comparison (Table 2).

... When the study area is narrowed down to the Pacific, the top 3 ranking variables are the same, with the net shortwave radiation takes the 1st place (Tab. 2). In both cases, the net shortwave and net longwave fluxes are ranked in top 3, which suggests that radiation budget is closely associated with SST inter-model differences. The pattern correlations of the 3 variables are higher within the Pacific than that at the global scale, indicating an increased robustness of the variable estimates.

... For simulation differences after changing the OGCM, latent heat flux holds the biggest share in pattern correlation with corresponding SST differences (Tab. 1), much bigger than that of net longwave flux at the second place and sensible heat flux at the third place. When the study area is confined within the Pacific, the three leading variables are the same but the sensible heat flux is ranked top (Tab. 2). Compared with the rankings in the AGCM case, it can be seen that the OGCM influences the simulation with different physical mechanisms than those by replacing AGCM. The effect of cloud radiation feedback is not prominent, while the physical processes associated with latent heat and sensible heat, including heat conduction, evaporation and convection, take a bigger share in the SST inter-model differences. SST differences can also lead to changes in longwave flux.’

*5. Please state significance test for pattern correlation, particularly how the effective degree of freedom is computed or decorrelation length scale used. Having a significance at 99.9% level for a correlation at 0.018 is hard to reckon with.*

Reply:

Thank you for pointing out the problem. We applied the Student-t test on the correlation coefficients, and the degree of freedom is taken as  $n-2$  ( $n$  is the number of grid points over the globe). Because the model data has been interpolated onto the  $1 \times 1$  grid, which makes the sample

points too large even without land points. The threshold value for 99.9% is 0.0174. We have added information of significance test in the table caption as follows:

‘Asterisk (\*) denotes the correlation coefficients passing the Student-t test above 99.9% confidence level. A larger number of grid points are involved that makes the threshold value relatively small. Numbers in boldface denotes the maximum absolute correlation value for the AGCM or OGCM replacement.’

*6. Page 6, line 13: Which year of external forcing (like CO<sub>2</sub>, greenhouse gases, aerosol, etc.) was used for the piControl run to obtain equilibrated state? Is this the same (similar) forcing as seen during observation period? If the piControl uses forcing that is quite different from observations (e.g. excess of 100ppm of CO<sub>2</sub>), then one would have to account for these differences when computing biases with respect to observations.*

Reply:

Thank you for these comments. The MPI-ESM experiment is configured based on the piControl run but not restricted to every parameter value. The external forcing is consistent with the present time climatology. The CO<sub>2</sub> value is set to default 353.9 ppm in the user manual. Other greenhouse gases like NO<sub>2</sub> also follows the default present time setting so that they are consistent with each other. The aerosol settings use the climatology compiled by S. Kinne without any complementation of volcanic aerosols. Model initialization is started from the climatology basic state recalculated with the AMIP run input data from 1981 to 2010. The piControl run data used in this study has the same time span as that of the reanalysis data from the year 1981 to 2010, which should be able to represent the model abilities in reproducing the climatology of the same period. Relevant information has been added to section 2.4.

## **‘2.4 Experimental Setup**

The control experiments conducted in this study are aimed to reproduce atmospheric and oceanic circulation characteristics of present time, and then to compare with each other in order to examine model performance improvements. The coupled control simulation is thus configured with reference to MPI-ESM piControl experiment settings. Model initialization is started from the climatology basic state recalculated with the AMIP run input data from 1981 to 2010. The CO<sub>2</sub> value is set to default 353.9 ppm in the user manual. Other greenhouse gases like NO<sub>2</sub> also follows the default present time setting so that they are consistent with each

other. The aerosol settings use the climatology compiled by S. Kinne without any complementation of volcanic aerosols....’

*7. page 12, line 7-8: Rather than inferring the relationship, the authors can easily compute the pattern correlation and quantify the correlation.*

Reply:

Thanks for your suggestion. We have calculated the pattern correlation between temperature biases and easterly anomalies in the three CGCM experiments. Because the easterly biases are most remarkable below 800hPa, the levels taken into consideration are between 800hPa and 1000hPa. The pattern correlation coefficients between temperature biases and zonal wind biases are 0.315 for the ECHAM6-NEMO3.6 experiment, 0.586 for the ECHAM5-NEMO3.6 experiment, and 0.411 for the MPI-ESM experiment. All of them have passed the 99% Student-t significance test. The ECHAM5-NEMO3.6 has the most significant easterly biases and holds the largest correlation coefficient, much bigger than that of the other two CGCMs, which indicates a strong correlation between the warmer temperature and stronger easterly anomalies. The analysis result has been added to section 5.2.

‘It can be seen from Figure 11 that cold temperature biases in the MPI-ESM transform into warm temperature biases in the ECHAM5-NEMO3.6 accompanied by increasing easterly anomalies in lower atmosphere, which suggests warmer temperature bias is correlated with stronger easterly anomalies. The pattern correlation between temperature biases and easterly anomalies in the three CGCM experiments has been calculated respectively. Because the easterly biases are most remarkable below 800hPa, the levels taken into consideration are between 800hPa and 1000hPa. The pattern correlation coefficients between temperature biases and zonal wind biases are 0.315 for the ECHAM6-NEMO3.6 experiment, 0.586 for the ECHAM5-NEMO3.6 experiment, and 0.411 for the MPI-ESM experiment. All of them have passed the 99% Student-t significance test. The ECHAM5-NEMO3.6 has the most significant easterly biases and holds the largest correlation coefficient, much bigger than that of the other two CGCMs, which confirms the correlation between the warmer temperature and stronger easterly anomalies.’

Minor comments

*1. page 6, line 11: please provide the actual websites*

Reply:

We have added the URLs of the websites for MPIOM and NEMO.



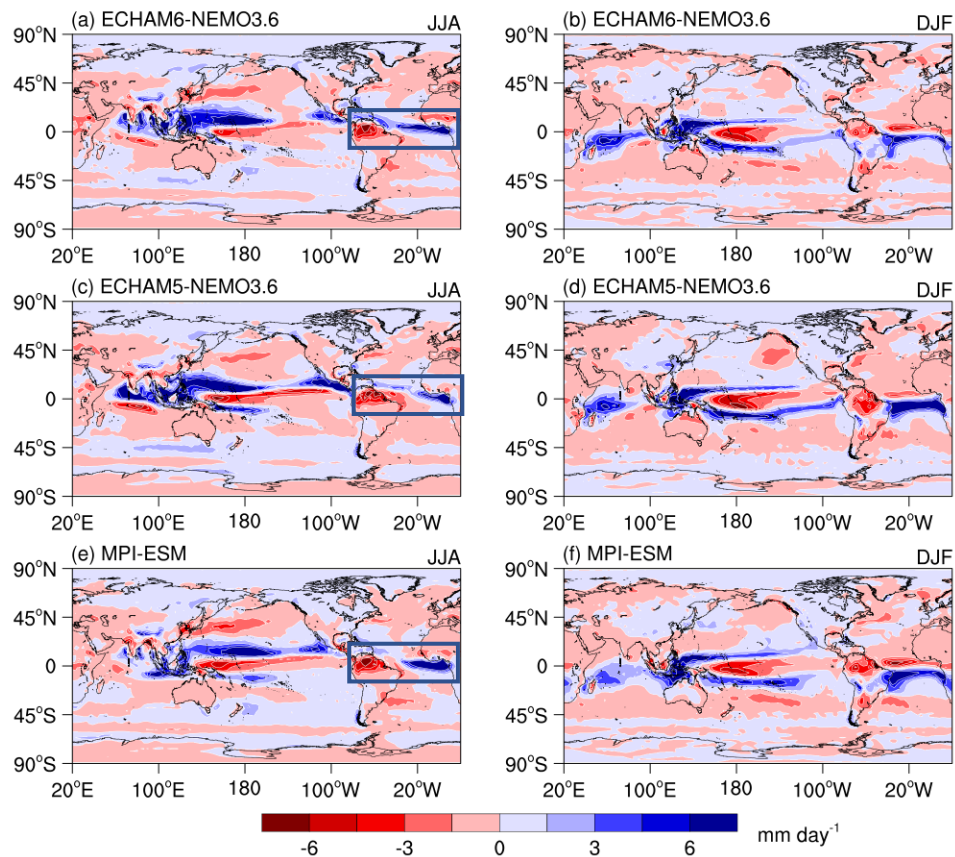
‘...namelist settings of MPIOM and NEMO3.6 still follow their own default settings for control run provided by their respective official websites (<http://www.mpimet.mpg.de/en/science/models/mpi-esm/mpiom> and <https://www.nemo-ocean.eu>).’

2. *page 8, line 31-32: Been staring at Fig. 4 but am not able to see “positive precipitation bias over the southern tropical Atlantic and negative bias on tropical South America”. Maybe you meant “negative bias over southern tropical Atlantic and positive bias on tropical South America”? Please indicate which specific figures you are referring to.*

Reply:

Thank you for pointing out the problem. We try to describe the variational trends of inter-model differences of precipitation that are consistent with previous research finding. It refers to the model biases in Figures 4a, 4c, 4e. We have corrected some imprecise expressions in the context. The focus areas have been outlined in the figure below.

‘...As suggested by Stevens et al. (2013), the extent of precipitation bias over the tropical Atlantic and South America seems related to each other, because the ECHAM5-NEMO3.6 shows less biases than those of the other two CGCMs with less blue contours over the tropical Atlantic and lighter red contours over Colombia and Venezuela (Fig. 4a, c, e) ....



3. page 8, line 32: “... seems related to each other, because the values and scales of biases in ... “Please explain, it is not obvious how this is the case.

Reply:

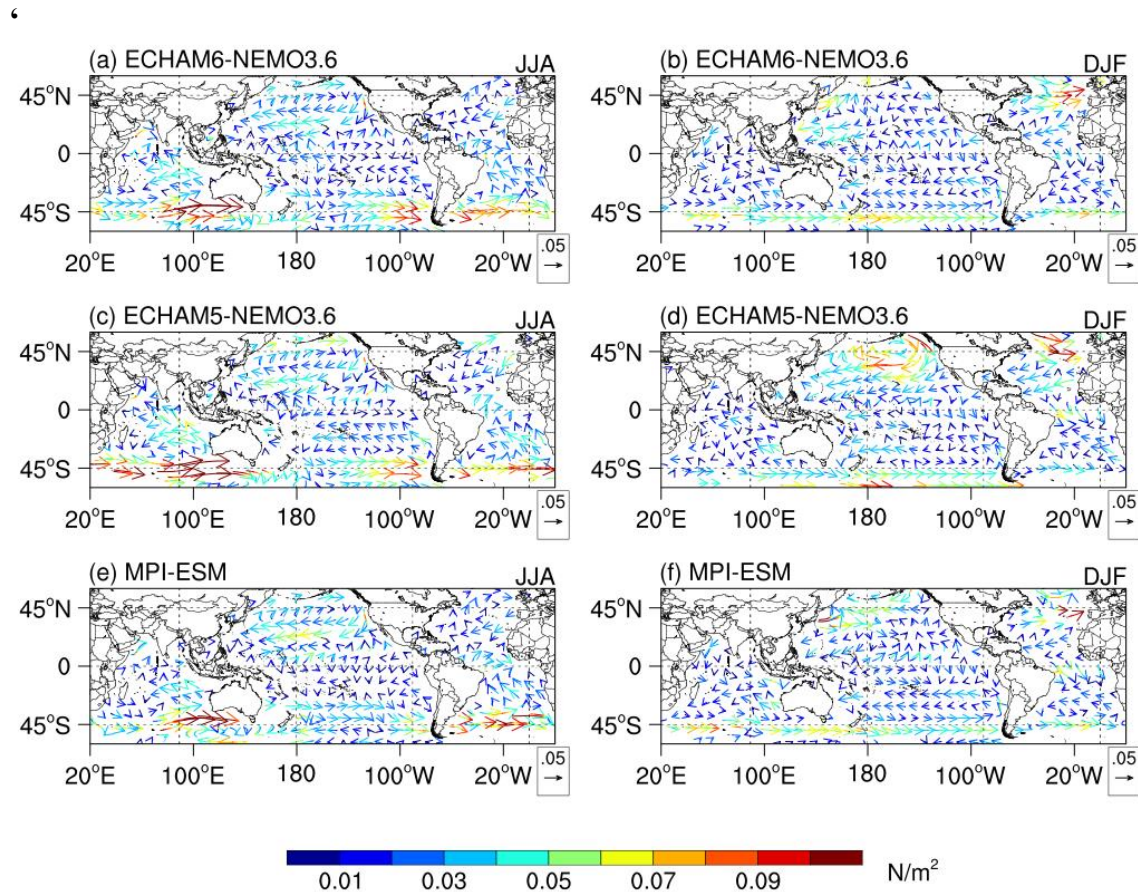
As response to the above question, the simulated precipitation exhibits inter-model differences within the boxes. The ECHAM5-NEMO3.6 shows less biases than those of the other two CGCMs, with less blue contours over the tropical Atlantic and lighter red contours over Colombia and Venezuela. The simultaneous reduction of precipitation biases coincides with the research findings of Stevens et al. (2013), who point out that the precipitation biases over the two areas are possibly related. We have further modified the paper content to clarify the issue.

‘...As suggested by Stevens et al. (2013), the extent of precipitation bias over the tropical Atlantic and South America seems related to each other, because the ECHAM5-NEMO3.6 shows less biases than those of the other two CGCMs with less blue contours over the tropical Atlantic and lighter red contours over Colombia and Venezuela (Fig. 4a, c, e) ....’

4. *Fig. 5: Please provide magnitude difference in surface wind stress in colour. This would make it easier for readers to see extent of anomalies.*

Reply:

We have shaded the vectors in Figure 5 as you suggest.



**Figure 5: The same as Fig. 3 but for simulated surface wind stress (unit: N/m<sup>2</sup>).’**

5. *page 9, line 15: Change “sea water to across” to “sea water across”*

Reply:

Thank you for pointing out the grammar mistake. We have revised the sentence.

‘...which can drive the underlying sea water across the date line in the other two CGCMs.’

6. *page 10, line 13: Change “it turns out that three CGCMs” to “it turns out that all three CGCMs”*

Reply:

Thank you for carefully reading. We have revised the sentence as you suggest.

‘...and it turns out that all three CGCMs suffer from a similar bias pattern in boreal summer ....’

7. *page 11, line 11: Change “,” to “.”*

Reply:

Thank you for pointing out the problem. We have corrected the mistake.

‘Bias in the MPI-ESM experiment manifests itself to a larger degree (Fig. 8c).’

8. *page 11, line 31: Change “In consistence” to “Consistent”*

Reply:

Thanks for your advice. We have revised the sentence as you suggest.

‘Consistent with surface wind stress bias (Fig. 5), ....’

9. *page 13, line 3: Change “it can be assumed” to “it would suggest”*

Reply:

Thanks for your suggestion. We have revised the sentence as you suggest.

‘Noting that latent heat flux has a correlation value close to that of net longwave radiation, it would suggest that cumulus convection changes sea surface winds and then surface evaporation, ....’

10. *page 15, line 22: None of the CGCMs used on this paper resolved mesoscale processes, so the subsentence “depending on mesoscale processes in the atmospheric and oceanic systems” cannot be concluded from this study and should thus be omitted.*

Reply:

Thank you for your advice. We have deleted the improper content from the sentence.

‘...model errors of each CGCM are caused by different physical mechanisms.’

11. *Please state units for Fig. 5, 6, 8, 9, 10, 11.*

Reply:

We have added the units in figure captions.

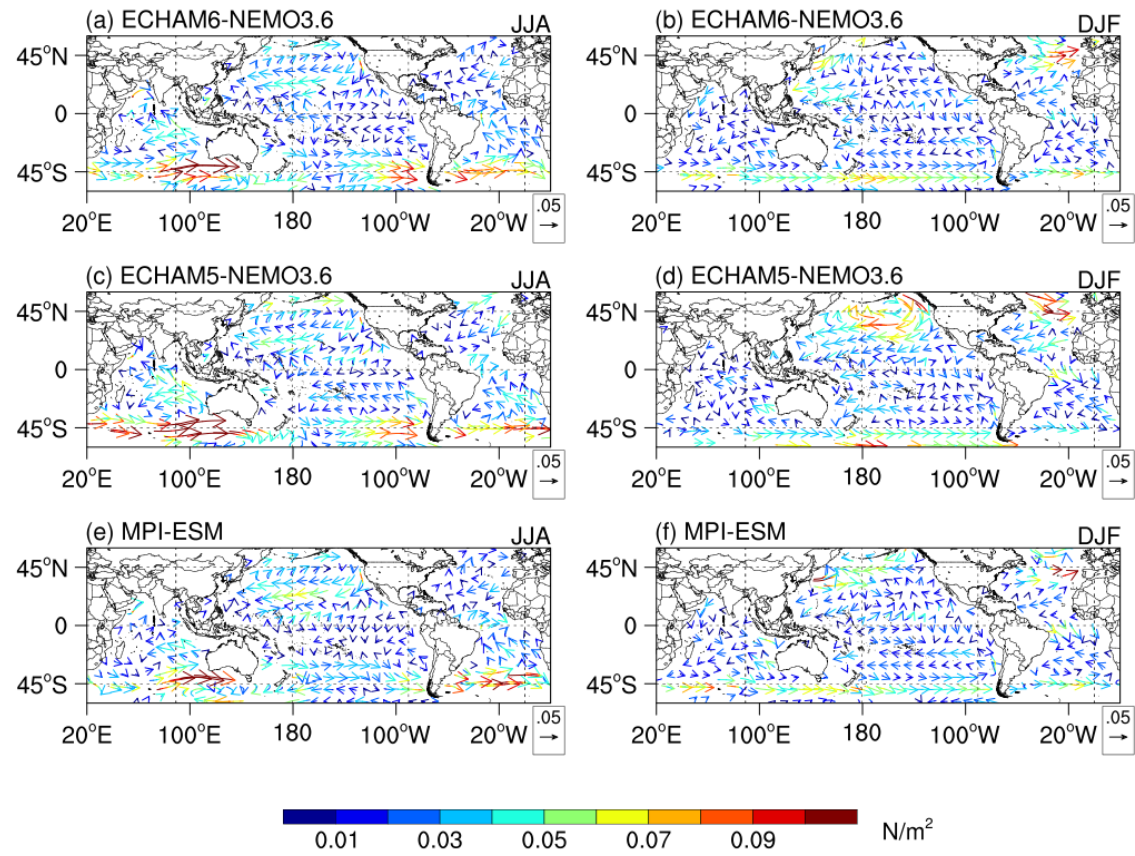


Figure 5: The same as Fig. 3 but for simulated surface wind stress (unit:  $\text{N/m}^2$ ).



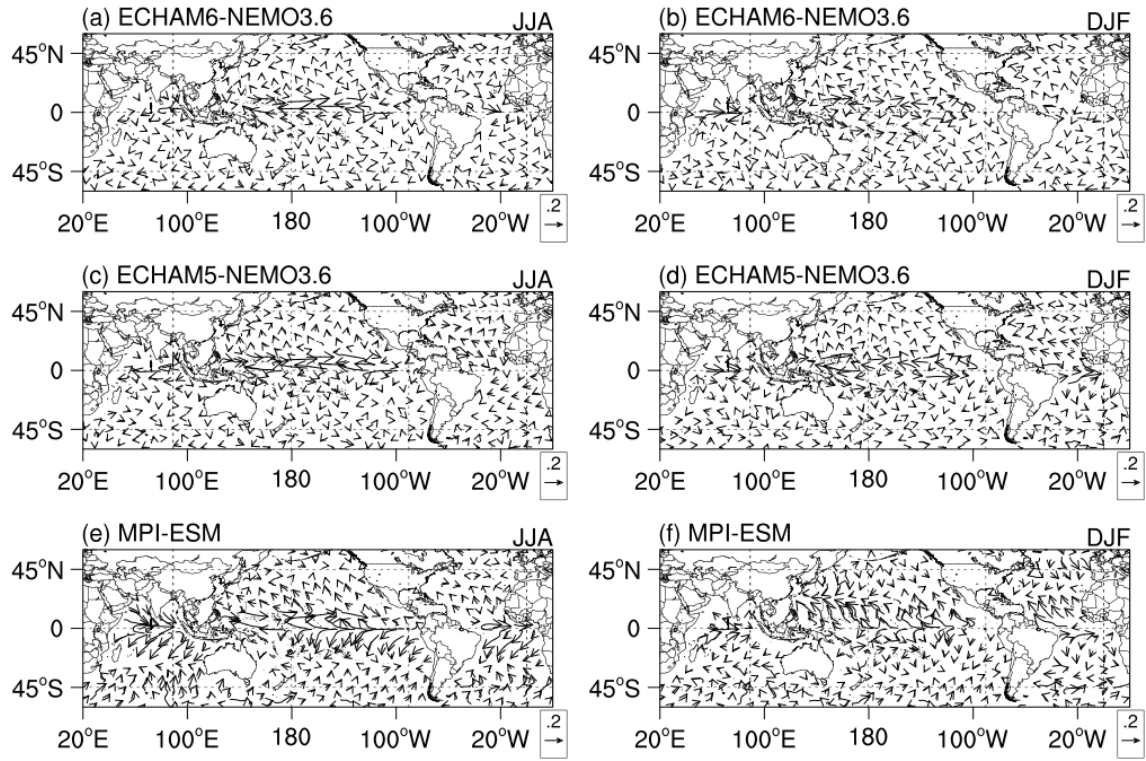
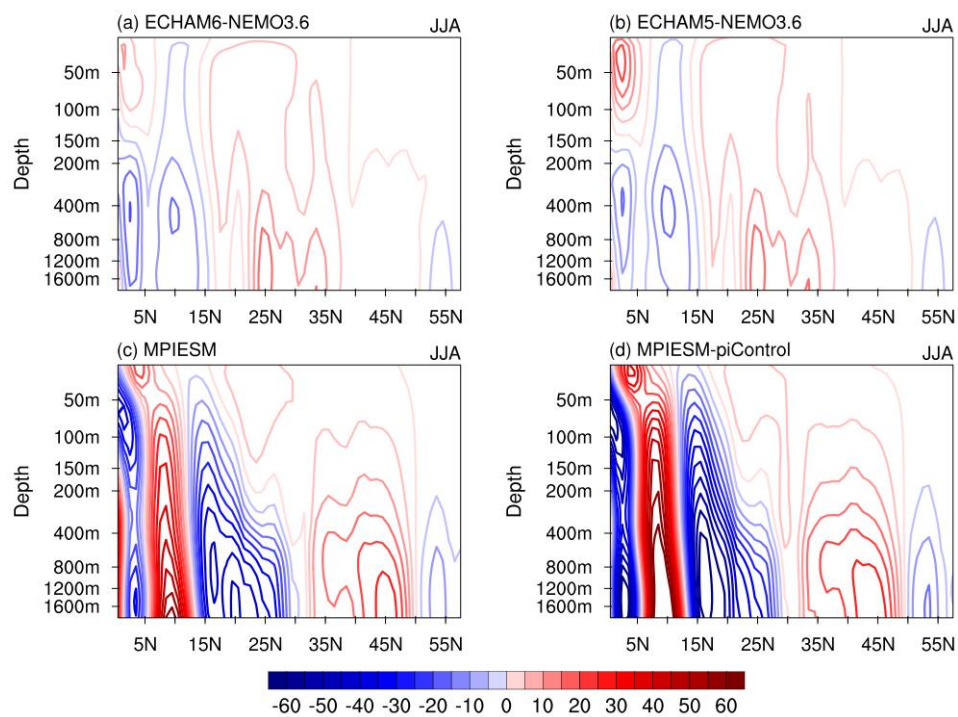
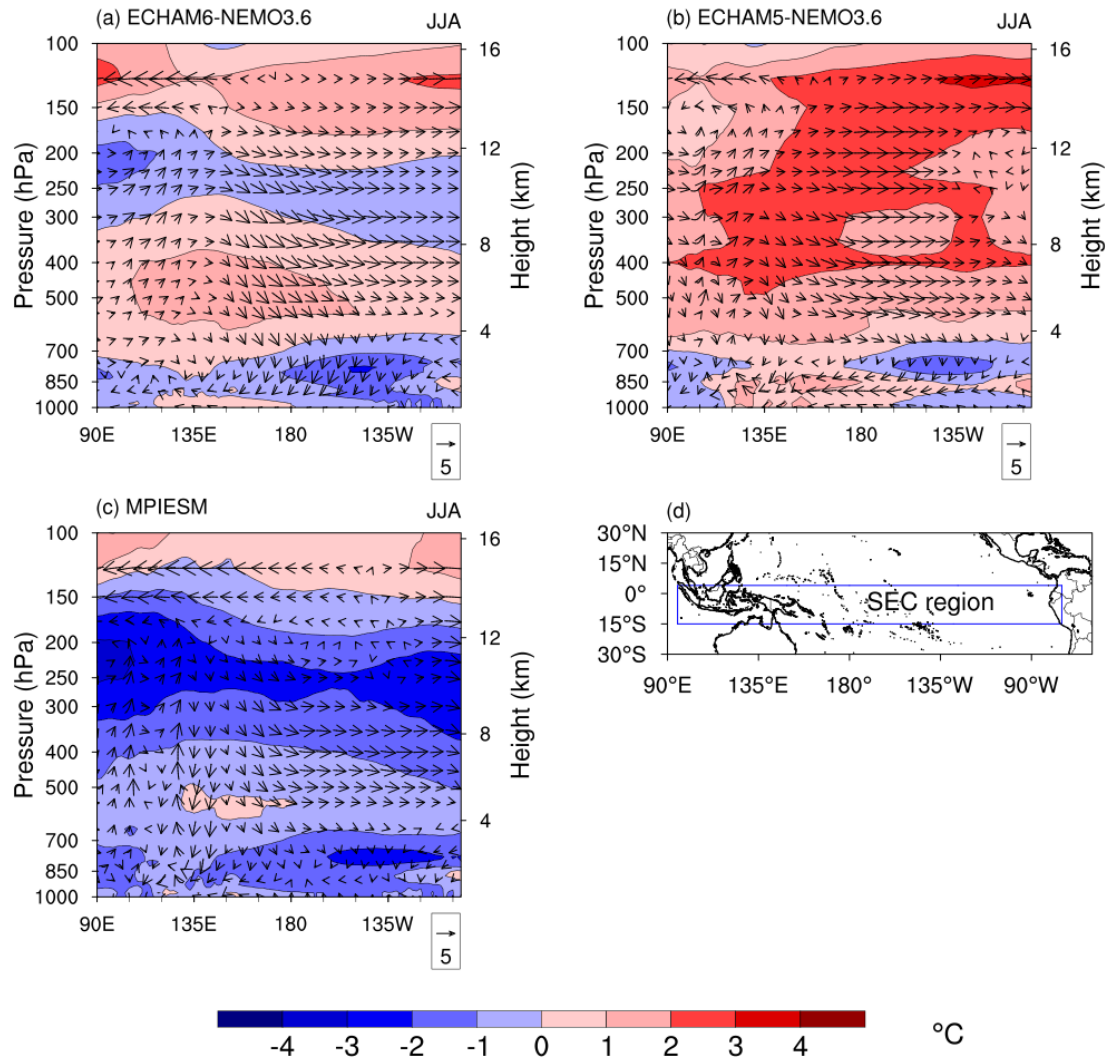


Figure 6: The same as Fig. 3 but for simulated surface currents (unit: m/s).

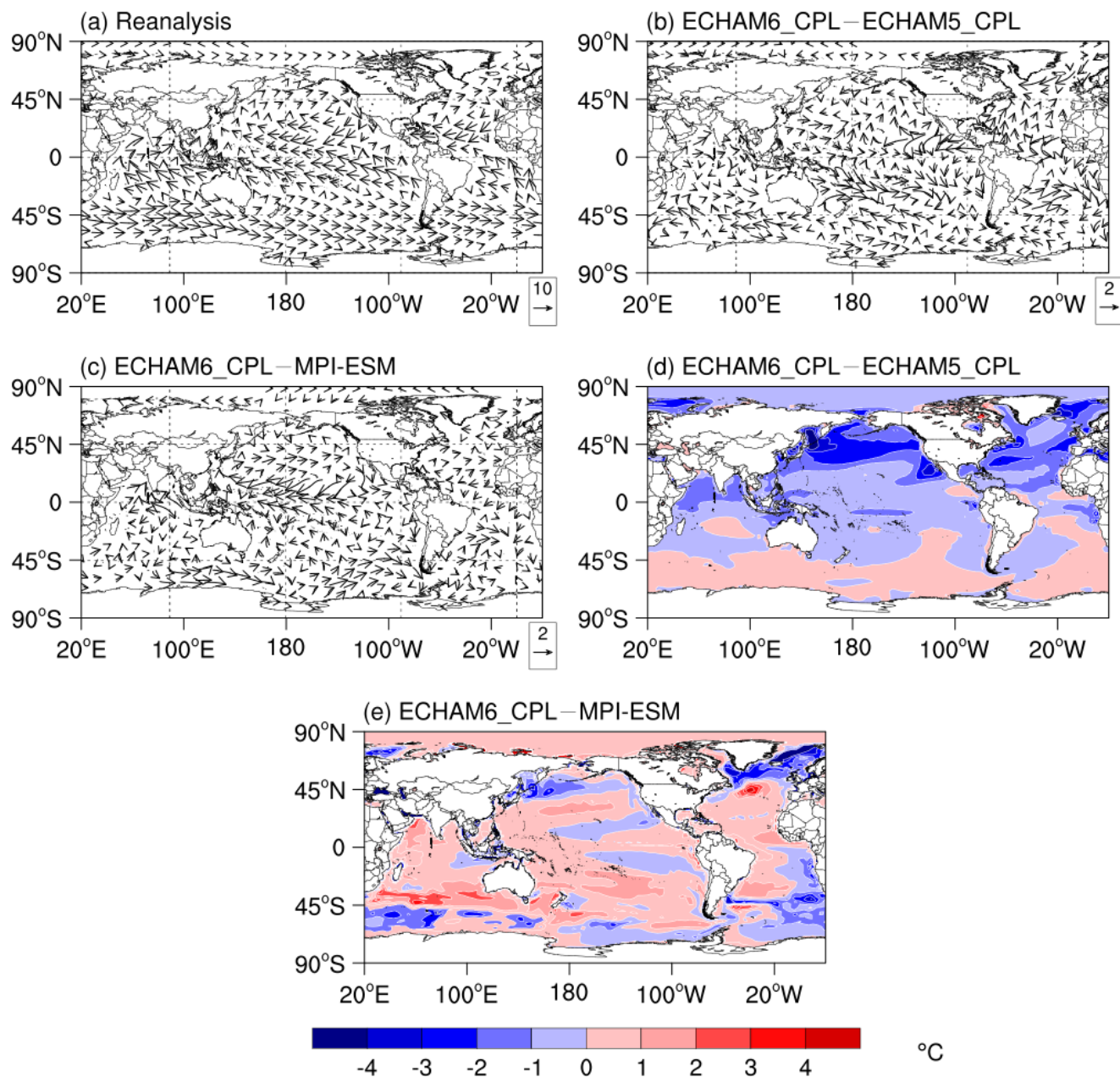


**Figure 8: Model biases of summer climatology of meridional overturning circulation simulation in North Pacific, (a) ECHAM6-NEMO3.6, (b) ECHAM5-NEMO3.6, (c) MPI-ESM, (d) MPI-ESM piControl, Units: Sv.**

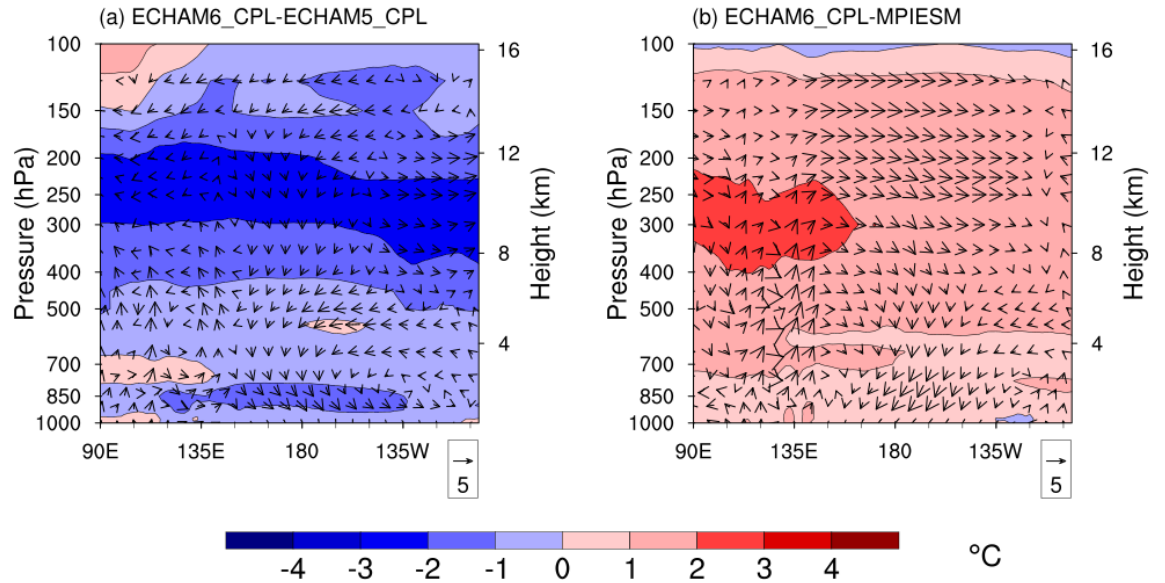


**Figure 9: Model biases of the vertical structure of atmospheric circulation (vector) and the temperature (contour) over SEC area in summer for (a) ECHAM6-NEMO3.6, (b) ECHAM5-NEMO3.6, (c) MPI-ESM. The SEC region is marked as the blue box in (d). Vector units: m/s, contour units: °C.**





**Figure 10: Summer climatology of 10m wind for (a) reanalysis data, (b) model differences between ECHAM6-NEMO3.6 and ECHAM5-NEMO3.6, (c) model differences between ECHAM6-NEMO3.6 and MPI-ESM. SST simulation differences for (d) between ECHAM6-NEMO3.6 and ECHAM5-NEMO3.6, and (e) between ECHAM6-NEMO3.6 and MPI-ESM. Vector units: m/s, contour units: °C.**



**Figure 11: Simulation differences in the vertical structure of atmospheric circulation and the temperature over the SEC area: (a) between ECHAM6-NEMO3.6 and ECHAM5-NEMO3.6, (b) between ECHAM6-NEMO3.6 and MPI-ESM. Vector units: m/s, contour unit: °C.**

,

The marked-up manuscript version is pasted below with yellow highlights on the changes:

# Simulation Improvements of ECHAM5-NEMO3.6 and ECHAM6-NEMO3.6 Coupled Models Compared to MPI-ESM and the Corresponding Physical Mechanisms

Shu Gui <sup>1</sup>, Ruowen Yang <sup>1</sup>, and Jie Cao <sup>1</sup>,

<sup>1</sup> Department of Atmospheric Sciences, Yunnan University, Kunming, 650091, China.

*Correspondence to:* Jie Cao (caoj@ynu.edu.cn) and Ruowen Yang (yangruowen@ynu.edu.cn)

**Abstract.** To improve the model simulation through decisive coupling mechanisms, rather than blindly updating the parameterization schemes, it is necessary to compare model performances between the CGCMs with the same atmospheric or oceanic component model. Therefore, two new CGCMs have been developed with the same oceanic component model, namely ECHAM5-NEMO3.6 and ECHAM6-NEMO3.6. The MPI-ESM that consists of ECHAM6 and MPIOM has also been employed. Experiments are carried out with the same settings in coupler and individual component model if applicable, and the new models show substantial improvements in the simulation of SST, precipitation and ocean currents. Further analysis has made it clear that the primary cause of SST biases in ECHAM5-NEMO3.6 and ECHAM6-NEMO3.6 can be attributed to the momentum field, while oceanic dynamics and surface radiation budget are accountable for more SST deviations in the MPI-ESM. Inter-model comparison between the coupled models with the same oceanic model suggests that cumulus convection is in the central part of simulation differences, which finally influence the SST through WES feedback mechanism. Whereas the OGCM replacement shows that latent heat of evaporation plays a predominant role in changing SST and surface radiation budget, and eventually bringing about variations in air temperature and atmospheric circulation. The mechanisms revealed in this study provide a new perspective of bias genesis during model coupling, which can be helpful for tuning other climate models towards a more realistic simulation.

## 1 Introduction

The physical processes sensitive to the air-sea interaction and its impact on climate variabilities have been studied for decades. Research findings suggest that dynamic air-sea coupling is important for tropical cyclone prediction regarding its intensity and rate of intensification (Sandery et al., 2010; Chen et al., 2013; Lin et al., 2018). The air-sea interaction has a substantial effect on precipitation response to ENSO teleconnections (Langenbrunner and Neelin, 2013), which is also associated with basic state climatology (Ham and Kug, 2015) in the Coupled Models Intercomparison Project phase 5 (CMIP5) models. The double intertropical convergence zone (ITCZ) problem of precipitation simulation in the coupled general circulation models (CGCMs) are closely linked with ocean-atmosphere feedbacks, including Bjerknes feedback, sea surface temperature (SST)-surface latent and surface shortwave flux feedback (Lin, 2007). The ocean-atmosphere interaction tends to amplify trade wind biases in general circulation models (GCMs) (Li and Xie, 2014), which is also responsible for excessive cold tongue simulation in equatorial Pacific. Besides traditional

understanding of coupling processes that account for SST and precipitation biases, recent studies have revealed other factors in the air-sea interaction that significantly contribute to the model bias pattern. Burls et al. (2017) find a quadratic relationship between extra-tropical Pacific albedo and equatorial SST bias, and Pham et al. (2017) suggested that the deep cycle of cold tongue turbulence can be affected by cloud cover and rain. The air-sea kinetic energy budget is found to be linked with surface gravity waves, where wave age and friction velocity affect in the ratio between kinetic energy from the winds and underlying surface currents (Fan and Hwang, 2017).

The dynamic mechanisms that play a major role in the biases propagation in the CGCMs are also investigated from a wide range of perspectives. Double ITCZ precipitation problem is usually associated with biases in radiation budget and surface winds (Lin, 2007). There is a close relationship between clouds and SST variation (Klein & Hartmann, 1993; Norris & Leovy, 1994), and changes of low clouds and shortwave radiation flux react on the SST and sea level pressure (SLP) (Norris et al., 1998; Mochizuki & Awaji, 2008; Bond & Cronin, 2008). Wu and Kinter III (2010) suggested that high-frequency changes in atmospheric circulation affects largely on the surface shortwave radiation, and hence its correlation with SST variability in the mid-latitude North Pacific. The weak simulation of Atlantic meridional overturning circulation (AMOC) is found to be responsible for sea surface temperature (SST) cold biases in the northern hemisphere in 22 CMIP5 climate models (Wang et al., 2014), which poses a great impact on the North Pacific through Northern Hemisphere annular mode (NAM) and wind-evaporation-SST (WES) feedback, combining SST biases in extratropical North Atlantic (ENA) and tropical North Atlantic respectively (Zhang & Zhao, 2015). The temperature bias can also be attributed to underestimation of water vapor amount (Liu et al., 2011), radiative and non-radiative processes (Ren et al., 2015), modelling of cloud-radiation feedback (Song et al., 2012) and gap winds (Sun & Yu, 2006). Wu and Liu (2003) suggested that the regional SST variations in central and eastern North Pacific are separately resulted from changes of Ekman advection and surface heat flux, both of which are affected by the subtropical ocean circulation, as a part of meridional overturning circulation in North Pacific (NPMOC). The NPMOC has been found to work as a bridge for mass and heat exchanges (McCreary and Yu, 1992; Liu et al., 1994), and is largely driven by sea surface wind stress and resulting in east-west sea level slope (Liu et al., 2011; Liu et al., 2013). However, there are larger discrepancies in the estimates of drag coefficients among different computational approaches, especially due to uncertainties in velocity measurements and removal of non-wind-driven currents.

Some efforts have been made to improve the SST simulation quality in a variety of CGCMs, for example, through changing zonal filtering and advection scheme (Xiao, 2006), modifying radiation and cumulus parameterization scheme (Bao et al., 2010), including frozen precipitating hydrometeors in cloud mass (Li et al., 2014), and decreasing relative humidity threshold for low cloud formation (Tang et al., 2016). However, these studies only focus on a limited range of processes that turn out to be important in the statistical analysis of model biases. Updates in the parameterization schemes bring in both improvements and setbacks in simulation. On behalf of oscillations in atmosphere-ocean coupling, the contribution of the individual

component model to the simulation of some key variables, regarding ENSO variability, extreme precipitation and hurricanes, and climate response to anthropogenic greenhouse gases, have not yet been made clear.

This paper studies the simulation improvements by new combinations of component models, analyses differences by respectively changing each component model, and determines the key physical processes behind the simulation deviations. To compare relative contributions of the individual component model in the CGCM, we need 2 CGCMs based on the same oceanic model but different atmospheric models, and another 2 CGCMs with the same atmospheric model but different oceanic models. For this purpose, two new CGCMs have been developed. One uses ECHAM5.4 as the atmospheric component model and NEMO3.6st as the oceanic component model, which is referred to as ECHAM5-NEMO3.6. The other uses ECHAM6.3 as the atmospheric component model and NEMO3.6st as the oceanic component model, and thus referred to as ECHAM6-NEMO3.6. The MPI-ESM developed by Max-Planck Institute for Meteorology, based on ECHAM6.3 for atmosphere and MPIOM for ocean, is also used in this study. To minimize simulation differences caused by model configurations, the three coupled models are set to the same coupling frequency of every 4 hours, and the same horizontal resolution T63 (192 longitudes×96 latitudes) for the atmospheric component model. The ocean models, namely MPIOM and NEMO3.6st, have different model structures and thus being used with their own default configurations. The content organization of this paper is as follows: A brief description of model frameworks and experiment setup configurations are presented in section 2. The reanalysis data sets used for model assessment and bias analysis are illustrated in section 3. Model evaluation and comparison among ECHAM5-NEMO3.6, ECHAM6-NEMO3.6 and MPI-ESM are presented in section 4. Further analysis on the cold tongue bias and opposite SST bias in North Pacific is presented in Section 5. Determination of key physical processes responsible for the differences in SST simulation after changing each component model is elaborated in Section 6. Summary and discussion part is in Section 7.

## **2 Model Description**

### **2.1 OGCM**

#### **2.1.1 NEMO**

NEMO model is a well renowned modelling system with high skills in global oceanic circulation simulation, which has been widely used for scientific research, weather forecast (Storkey et al., 2014; Megann et al., 2014), and reanalysis data assimilation (Mogensen et al., 2012a, b). Designed to serve as a flexible tool for ocean and sea ice studies, NEMO manifests good usability interacting with other ACGMs (Gualdi et al., 2003; Luo et al., 2005; Park et al., 2009; Dunlap et al., 2014; Huang et al., 2014). The NEMO stable version 3.6 has been employed in this study, whose ocean component is configured to calculate primitive equations on the ORCA2 grid, a tripolar grid of 182 (longitude)×149 (latitude) curvilinear orthogonal mesh in horizontal direction and 31 vertical levels unevenly distributed on partial step z coordinate in current research. Turbulent kinetic energy (TKE) closure scheme has been chosen for vertical mixing with enhanced vertical diffusion for convective processes. The Louvain-la-Neuve sea-ice model (LIM3), originally developed by

Fichefet and Morales-Maqueda (1997), has been incorporated in NEMO3.6 to represent the sub-grid-scale dynamics and their impact on sea ice thickness and ice-ocean salt exchanges. Main differences between LIM3 and other ice models are related to the physical parameterization of open boundary conditions and sea-ice interactions, with the C-grid formulation of elastic-viscous-plastic rheology (Bouillon et al., 2013).

### **2.1.2 MPIOM**

MPIOM model is formulated on Arakawa-C grid for horizontal dimension and z-grid for vertical dimension, using the hydrostatic and Boussinesq approximations in the model dynamic equations (Jungclauss et al., 2006; Jungclauss et al., 2013). Vertical mixing and diffusion are parameterized following Pacanowski and Philander (1981), with a diffusion coefficient varies with grid spacing (Redi, 1982). Sea-ice model is included in MPIOM where sea-ice thickness is modulated by turbulent atmospheric fluxes and oceanic heat transport (Wolff et al., 1997; Marsland et al., 2003; Notz et al., 2013). The model is configured with 40 unevenly spaced vertical levels on the GR1.5 grid, a conformal mapping grid of 256 (longitude)  $\times$  220 (latitude) in the horizontal making horizontal resolution approximately 1.5°.

## **2.2 AGCM**

The ECHAM atmospheric model developed by the Max Planck Institute for Meteorology has been used in many studies since its first version (ECHAM1) branched from the cycle 17 operational model at Medium Range Weather Forecasts (ECMWF) (Roeckner et al., 1989; Simmons et al., 1989). Incorporation of new features along the course of model development gradually makes ECHAM capable of reproducing meticulous characteristics in the weather system, including those in cumulus convection, moisture transport, radiation and land-surface processes (Roeckner et al., 1996, 2003; Raddatz et al., 2007; Brovkin et al., 2009). ECHAM has also been employed in the coupled earth modelling system, from coupling with large - scale geostrophic ocean model (LSG) (Maier - Reimer et al., 1993) to the latest version of Earth system model MPI-ESM (Baehr et al., 2015). The ECHAM model consists of a dry spectral dynamic core, a set of parameterization schemes dealing with solar irradiance, moist convection, land-surface properties, etc. The versions in use for this paper are ECHAM5.4 (Roeckner et al., 2003) and ECHAM6.3 (Stevens et al., 2013). Major updates from ECHAM5 to ECHAM6 include improved representation of shortwave spectrum, a new aerosol parameterization scheme, middle atmosphere and surface albedo descriptions are also enhanced. Despite the new implementations in ECHAM6, the AGCMs are set to the same configuration if applicable, to minimize the differences caused by updates of physical parameterization schemes.

## **2.3 CGCM**

### **2.3.1 ECHAM5-NEMO3.6**

The schematic structure of ECHAM5-NEMO3.6 is shown in Fig. 1a. Overall, the ECHAM5-NEMO3.6 consists of the atmospheric model ECHAM5.4, oceanic and sea ice model NEMO3.6st (the stable version of NEMO3.6), and the coupler Ocean Atmosphere Sea Ice Soil (OASIS3) (Valcke, 2013). Although ECHAM5

(Roeckner et al., 2003, 2006) is an older version of the atmospheric model developed by Max Planck Institute of Meteorology compared with ECHAM6 (Stevens et al., 2013), it was employed by the previous coupled atmosphere ocean model ECHAM5/MPI-OM (Jungclaus et al., 2006), which has been well tested and reassured of accurate surface flux transfer between the oceanic and atmospheric component models. Based on the interchange coupling structure of the ECHAM5/MPI-OM, seventeen variables are passed from ECHAM5 to NEMO-3.6st through OASIS3 coupler, including solar radiation, non-solar heat flux and its derivative with respect to temperature, zonal and meridional wind stress, evaporation minus precipitation, and sublimation. The meridional and zonal wind stress vectors are passed to the ORCA2 U and V grid of NEMO-3.6st, while other variables are passed to the T grid. Both ocean and ice regions are considered in the coupling processes. In the opposite direction, six variables comprised of the SST, sea ice temperature, sea ice fraction, sea ice albedo and surface ocean currents are transferred from the ocean model to the atmosphere model. Bilinear interpolation method is used in the exchanges of physical variables between ECHAM5 gaussian grid and NEMO-3.6st ORCA2 grid. The time steps for ocean and atmosphere models are both set to 1200 seconds, and the coupling frequency is 4 hours once (every 12-time steps). At present, the component models are integrated with the default parameter settings suggested in the user manual, model retuning will be scheduled in further research.

### **2.3.2 ECHAM6-NEMO3.6**

Following the coupling framework of MPI-ESM (the update version of ECHAM5/MPI-OM, version number MPI-ESM-1.2.00p4), ECHAM6-NEMO3.6 has been developed with the same atmospheric component model ECHAM-6.3, coupled with NEMO 3.6 stable version through OASIS3-MCT (Craig et al., 2017) (Fig. 1b). Model retuning of the ECHAM-6.3 and NEMO3.6 is left for further studies due to time limitation and high computational cost. Namelist settings of the ECHAM-6.3 and the coupler OASIS3-MCT are brought into correspondence with those in ECHAM5-NEMO3.6 to the utmost, for example, with the same horizontal resolution T63 on gaussian grid and the same parameterization settings for greenhouse gases. The oceanic component model NEMO3.6 still uses the same configuration as that in ECHAM5-NEMO3.6. Coupling variables of the ECHAM6-NEMO3.6 are the same as those in ECHAM5-NEMO3.6. Details in experiment configuration are elaborated in section 2.4.

### **2.3.3 MPI-ESM**

The MPI-ESM is comprised of the atmospheric general circulation model ECHAM6 and the oceanic circulation model MPIOM (Jungclaus et al., 2013; Stevens et al., 2013). It has been continuously developed at Max Planck Institute for Meteorology and has been successfully applied on a broad range of studies, including volcano studies (Zanchettin et al. 2013; Zhang et al. 2013), anthropogenic land cover change (Reick et al., 2013; Brovkin et al., 2013), circulation feedback sensitive to Intertropical Convergence Zone (ITCZ), and double ITCZ precipitation (Mobis and Stevens, 2012). MPI-ESM has been used in CMIP5 and is employed in the upcoming CMIP6. Due to high computational cost, the low-resolution version MPI-ESM-

LR is used in this study (version number MPI-ESM-1.2.00p4), with ECHAM6 running at T63L47 resolution (horizontal resolution about  $1.875^\circ$  and 47 vertical levels). Experimental settings are migrated from piControl default configuration, and then adjusted to being in consistence with those of ECHAM5-NEMO3.6 and ECHAM6-NEMO3.6 experiments. The major differences from default piControl experiment are the increased coupling frequency (4 hours once), and climatology recalculated from the year 1981 to 2010 to serve as the model input files (e.g. aerosol properties, ozone mole fractions and land use transitions).

## 2.4 Experimental Setup

The control experiments conducted in this study are aimed to reproduce atmospheric and oceanic circulation characteristics of present time, and then to compare with each other in order to examine model performance improvements. The coupled control simulation is thus configured with reference to MPI-ESM piControl experiment settings. The CO<sub>2</sub> value is set to default 353.9 ppm in the user manual. Other greenhouse gases like NO<sub>2</sub> also follows the default present time setting so that they are consistent with each other. The aerosol settings use the climatology compiled by S. Kinne without any complementation of volcanic aerosols. Model initialization is started from the climatology basic state recalculated with the AMIP run input data from 1981 to 2010. The atmospheric component models (ECHAM-5.4 and ECHAM-6.3) used in this study are running on T63 gaussian grid (approximately equivalent to  $1.875^\circ \times 1.875^\circ$  on average), with the same coupling frequency of 4 hours to exchange momentum and heat fluxes with the ocean component model. Physical parameterization schemes relative to solar irradiance, aerosol optical properties, cumulus convection and strong stratospheric damping are maintained the same among ECHAM5-NEMO3.6, ECHAM6-NEMO3.6, and MPI-ESM experiments. Since the timestep length of ECHAM-6.3 in coupled mode is suggested to be 450 seconds according to its user manual, only the ECHAM5-NEMO3.6 experiment uses 1200 seconds as timestep length for the atmospheric model ECHAM-5.4. The oceanic component models (MPIOM and NEMO3.6) possess different model structures and mapping technologies, making it impossible to directly migrate physical parameterization settings from one to the other. In this regard, namelist settings of MPIOM and NEMO3.6 still follow their own default settings for control run provided by their respective official websites (<http://www.mpimet.mpg.de/en/science/models/mpi-esm/mpiom> and <https://www.nemo-ocean.eu>). Three coupled experiments, namely ECHAM5-NEMO3.6, ECHAM6-NEMO3.6, and MPI-ESM experiments, have been conducted for 200-year realizations including spin-up runs. Model results spanning the last 100 years (model year 101 to 200) should be well-equilibrated, and thus they are used to compute simulation climatology for the inter-model comparison analysis.

## 3. Reanalysis Data

The assessment of model performance with reference to each CGCM employs the monthly data from the Hadley Center (HadISST) (Rayner et al., 2003). Model precipitation evaluation uses reanalysis data of the Global Precipitation Climatology Project (GPCP) (Adler et al., 2003). Evaluation of the mean sea level pressure, zonal and meridional winds at 10m height, cloud cover and surface temperature use the ERA-



Interim monthly reanalysis data (Simmons et al., 2006). Due to the limited time span of the ERA-Interim, the ERA-20c reanalysis has been used instead for the evaluation of ENSO and SAM variability. For Surface wind stress data from the Scatterometer Climatology of Ocean Winds (SCOW) including QuikSCAT measurements (Risien & Chelton, 2008) has been chosen for its more advanced stress-measuring instrument and better sampling, which has been widely used in researches regarding oceanic circulation and dynamical processes (Kanzow et al., 2010; Roquet et al., 2011; Johnson et al., 2012) and evaluation of CGCMs and reanalysis data (Xue et al., 2011; Lee et al., 2013). To characterize the changes in ocean circulation associated with the SST bias, the SODA reanalysis data (Carton & Giese, 2008) has been used following the massive researches on ocean variability and mechanisms (Dewitte et al., 2009; Tett et al., 2014; Drenkard & Karnauskas, 2014). Finally, surface net radiation flux from CERES EBAF-Surface Ed4.0 (Kato et al., 2013) is employed for its higher accuracy by using more accurate cloud data to calculate solar radiation at the Earth's surface (Wild et al., 2013; Wild et al., 2015; Zhang et al., 2015), which is better than popular reanalysis data sets including NCEP-DOE, MERRA and ERA-Interim (Zhang et al., 2016). Because the local mesh refinements of ORCA2 grid in NEMO3.6 make meridional resolution finer in tropics (about  $0.5^\circ$ ) than that of higher latitudes (about  $2^\circ$ ), the SST model data has been remapped onto  $1^\circ \times 1^\circ$  as a trade-off. The precipitation from ECHAM-5.4 and ECHAM-6.3 (T63 resolution) has been re-gridded to the GPCP data grid ( $2.5^\circ \times 2.5^\circ$ ). Other physical quantities are interpolated onto  $1^\circ \times 1^\circ$  grid for both model and reanalysis data. As SODA3.3 reanalysis data has more vertical levels than NEMO model output, especially in the upper ocean, the OGCM velocity output has been interpolated onto SODA3.3 vertical levels to better compare the vertical circulation. The time period of reanalysis climatology spans from 1981 to 2010, except for SCOW with 122 months from September 1999 to October 2009 and CERES EBAF from 2000 to 2015. The differences between model and reanalysis data are defined as the anomalous fields with simulation minus reanalysis counterpart.

## 4. Model Evaluation and Simulation Improvements

### 4.1 Overall Performance

The simulated climatology of 9 physical variables, which are sensitive to coupling processes of each CGCM on the global domain, has been evaluated with Taylor's diagram analysis (Fig. 2). Summer and winter seasons are chosen because they represent the opposite extreme conditions in the course of model integration. The Taylor diagram shows satisfactory model performances for most physical quantities' simulation in the opposite seasons, except for meridional currents of MPI-ESM that lie beyond 2 times the standard deviation. It is noteworthy that these errors are not caused by changes in the MPI-ESM experiment settings, which is subject to discussion later by comparing with piControl experiment results in Section 5. The best simulations among the three CGCMs in summer are topped by 2m temperature of ECHAM5-NEMO3.6 with the maximum pattern correlation coefficient 0.995, followed by SST simulation in ECHAM6-NEMO3.6 and total radiation flux in MPI-ESM. For the rest of variables other than ocean surface currents, the standard

deviations lie between 0.953 and 1.336, which represent an unremarkable deviation in spatial variability of model simulation compared to reanalysis counterparts. The simulation differences for each variable in terms of root-mean-square deviations (RMSD) are below 0.840, with smaller values accompanied by higher correlation coefficients. The distributions of these statistical values in winter season are a little bit different, marked by 0.996 for SST simulation in ECHAM6-NEMO3.6 as the top, and followed by 2m temperature also in ECHAM6-NEMO3.6 and total radiation flux in MPI-ESM. Standard deviations of all simulated variables except for ocean currents are between 0.865 and 1.218, and RMSD values are within 0.765, both of which show a little improvement than the summer case. Precipitation simulation in MPI-ESM experiment for both seasons resembles that at the annual scale for MPI-ESM-LR experiments in Stevens et al. (2013).

The model performances shown in Fig. 2 are reasonably good compared with previous studies evaluating CMIP5 models (Sheffield et al., 2013), with simulation improvements relative to surface temperature and winds in ECHAM6-NEMO3.6 and ECHAM5-NEMO3.6. Improvements in climatology simulation are resulted from changing component model, for example, 2m temperature simulation in ECHAM6-NEMO3.6 is better than that in MPI-ESM, because the ocean model MPIOM in MPI-ESM has been replaced by NEMO3.6st to make ECHAM6-NEMO3.6. It arouses interests in how component model replacement affects the simulation in detail. Therefore, a few key variables in the course of coupling are selected to do inter-model comparison.

## 4.2 Sea Surface Temperature

The SST properties are among the most important factors that decides air-sea fluxes in a coupled system. Model biases of SST seasonal climatology in summer (June-July-August, JJA) and winter (December-January-February, DJF) are presented in Fig. 3. The differences between model simulations and reanalysis counterparts are usually less than 2°C except for polar areas. Reanalysis data itself is responsible for large biases near polar oceans, because there are also large biases in the previous research by Huang et al. (2014) using the same reanalysis dataset. Taken into account SST biases in both summer and winter, it can be seen that the spatial distribution of SST bias in the MPI-ESM experiment (Fig. 3e, f) resembles main characteristics of the annual SST bias in Jungclaus et al. (2013). In boreal summer, three CGCMs exhibits excessive cold tongue simulation with the minimum extent in ECHAM5-NEMO3.6 (Fig. 3c, d) and the maximum extent in MPI-ESM (Fig. 3e, f). There are substantial cold biases in western North Pacific and North Atlantic in ECHAM6-NEMO3.6 and MPI-ESM experiments. However, the ECHAM5-NEMO3.6 exhibits warm SST biases instead of cold biases in North Pacific, and no remarkable SST bias is found North Atlantic. The opposite SST biases in North Pacific may imply different physical mechanisms that take effects on the coupling processes, which will be discussed later. In boreal winter, the ECHAM6-NEMO3.6 only shows remarkably warm SST biases over 2 °C on southern tropical Atlantic, while other two CGCMs still present significant SST biases in other parts of the globe. The ECHAM5-NEMO3.6 features warm SST biases over the southern tropical ocean, with a large area of warm SST bias over the eastern Pacific. In contrast, MPI-ESM shows significantly cold SST biases over northern subtropical Pacific, mid-latitude southern

Indian Ocean, and mid-latitude northern Atlantic. The Southern Ocean SST nevertheless takes on warm biases in the MPI-ESM experiment.

Model improvements of the SST simulation can be summarized as follows: ECHAM5-NEMO3.6 best decreases the scope and intensity of excessive cold tongue and presents no remarkable biases over North Atlantic in boreal summer, and ECHAM6-NEMO3.6 reproduces the best winter climatology with only tropical South Atlantic suffering from more than 2 °C bias. Their SST simulation qualities are substantially improved compared with that of MPI-ESM, both from Taylor's diagram results (Fig. 2) and model deviations (Fig. 3).

### 4.3 Precipitation

Inter-model comparison of precipitation climatology simulation is presented in Fig. 4. Precipitation bias pattern of MPI-ESM experiment (Fig. 4e, f) generally resembles that of MPI-ESM-HR experiment at the annual scale in Stevens et al. (2013). ECHAM5-NEMO3.6 and ECHAM6-NEMO3.6 show similar bias patterns as that of MPI-ESM with variations in detail. The main errors in precipitation simulation are inadequate rainfall along the equatorial areas, flanked by excessive precipitation in subtropics of both hemispheres. The double ITCZ problem in summer is still remarkable in both MPI-ESM and ECHAM5-NEMO3.6, but it is ameliorated in ECHAM6-NEMO3.6 over South Pacific Intertropical Convergence Zone (SPCZ). As suggested by Stevens et al. (2013), the extent of precipitation bias over the tropical Atlantic and South America seems related to each other, because the ECHAM5-NEMO3.6 shows less biases than those of the other two CGCMs with less blue contours over the tropical Atlantic and lighter red contours over Colombia and Venezuela (Fig. 4a, c, e). Inter-model simulation differences may thus imply an important role of atmospheric dynamics simulation, for example storm-track and ITCZ, in reducing precipitation biases. Rainfall biases in winter exhibit larger dry biases in tropical Pacific, also flanked by wet biases in subtropics with significantly overestimation in South Africa and South Indian Ocean.

Model improvements of the precipitation simulation can be summarized as follows: The ECHAM6-NEMO3.6 reproduces the best overall precipitation climatology with the highest pattern correlation coefficient (Fig. 2). ECHAM6-NEMO3.6 model ameliorates double ITCZ problem over SPCZ in boreal summer, while ECHAM5-NEMO3.6 best decreases the wet bias over tropical Atlantic.

### 4.4 Surface Wind Stress

Wind stress biases in boreal summer (Fig. 5a, c, e) exhibit similar bias patterns among the three CGCMs, with anomalous anticyclonic circulations over North Pacific, South Atlantic and South Indian Ocean. There are prevailing westerly anomalies over the Southern Ocean that reinforce the anticyclones in mid-latitudes of southern hemisphere. The major difference among the three CGCMs is in tropical Pacific, where MPI-ESM features southerly (northerly) biases but ECHAM5-NEMO3.6 and ECHAM6-NEMO3.6 presents east (west) oriented biases north (south) of the equator. It implies that the excessive cold tongue simulation in MPI-ESM has little to do with surface wind stress anomalies, which can drive the underlying sea water across

the date line in the other two CGCMs. The anticyclonic bias over North Pacific can affect the meridional overturning circulation and subsequent heat transport. Therefore, the opposite SST bias in North Pacific between ECHAM5-NEMO3.6 and ECHAM6-NEMO3.6 can be partly attributed to differences in wind stress and radiation simulation. In boreal winter, only westerly anomalies are still prominent over the Southern Ocean among three CGCMs, whereas model biases in other parts of the world display arbitrary patterns. ECHAM5-NEMO3.6 still suffers from anticyclonic biases over North Pacific in winter, which has been largely diminished in ECHAM6-NEMO3.6.

The three CGCMs possess similar bias patterns at the global scale, without obvious improvements in wind stress simulation after changing component models. However, it is worth noting that the common SST bias over tropical Pacific may be ascribed to different physical mechanisms for MPI-ESM, because there only exist southward oriented wind biases.

#### 4.5 Surface Ocean Current

The ocean current biases are mainly located in tropical areas (Fig. 6), in predominantly zonal directions for ECHAM5-NEMO3.6 and ECHAM6-NEMO3.6, but meridionally distributed for MPI-ESM. Anomalous currents move westward south of the equator but turn to the opposite direction in northern tropics during boreal summer, with more intensified currents in ECHAM5-NEMO3.6 than those in ECHAM6-NEMO3.6 (Fig. 6a, c). Whereas the MPI-ESM features southward (northward) tilted biases south (north) of the equator to a larger degree than the other two CGCMs. The direction of ocean current biases generally agrees with that of wind stress biases, where poleward deflection can be attributed to Coriolis effects. Since the poleward motion is too strong in the MPI-ESM experiment, the ocean currents in subtropical North Pacific even turn to the east. There are northward current anomalies in North Pacific for both MPI-ESM and ECHAM5-NEMO3.6, favouring more heat transport from subtropics to higher latitudes. Yet colder SST biases still exist over large maritime space in MPI-ESM experiment, which suggests an investigation on radiation budget and the meridional overturning currents that provide a full picture of most relevant oceanic processes. The SST biases are also attributed to different projection grids of NEMO3.6 and MPIOM, which however is beyond the scope of this paper. Biases in winter season are diminished to some degrees in tropical oceans, with little amplification of biases outside tropics in ECHAM5-NEMO3.6 and ECHAM6-NEMO3.6 experiments. But the MPI-ESM comes up with prominent northward biases in subtropical North Pacific (Fig. 6f). It may help to explain warm SST bias around Sea of Japan (Fig. 3f) and cold SST bias in the subtropical ocean.

Although the structure of model bias differs between MPI-ESM and the other two CGCMs, the bias amplitude has been substantially reduced especially in tropics by using NEMO3.6st as the oceanic component model. The marked deficit in the simulation of ocean currents in tropical Pacific may be related to more in situ SST biases of the MPI-ESM than ECHAM5-NEMO3.6 and ECHAM6-NEMO3.6.

#### 4.6 Total Radiation

Figure 7 shows the total radiation biases in each CGCM, and it turns out that all three CGCMs suffer from a similar bias pattern in boreal summer, with overheated areas around the Pacific warm pool and insufficient total irradiance over the tropical eastern Pacific, subtropical southern Indian Ocean and North Atlantic. The MPI-ESM and ECHAM6-NEMO3.6 ameliorate the bias amplitude in ECHAM5-NEMO3.6, which is likely due to parameterization updates in the middle atmosphere that result in a more realistic simulation of cloud radiation forcing. Underestimation of surface net irradiance over the eastern Pacific leads to cooler SST, which may exaggerate the cold SST bias in tropical Pacific by anomalous ocean currents. Since the bias amplitude in ECHAM5-NEMO3.6 is the biggest, which is consistent with statistical analysis in Taylor's diagram (Fig. 2), the consequent SST bias should be larger than at least one of the other two CGCMs that either has the minimum deviation (MPI-ESM) or has the same oceanic component model (ECHAM6-NEMO3.6). However, the SST bias ranking is completely the opposite, especially around tropical Pacific (Fig. 3a, c, e). It implies that oceanic dynamics may be the major source for cold SST bias in the MPI-ESM, according to assessment results of wind stress, surface currents and radiation, whereas momentum fluxes may be the predominant factor for the other two CGCMs. Biases in boreal winter transform into an approximately zonal-band structure, including overestimation in tropical areas and underestimation in mid-latitudes (Fig. 7b, d, f). Cold biases still remain in limited areas of tropical Pacific and Atlantic, while considerable deviations occupy most part of subtropical and mid-latitude areas for ECHAM5-NEMO3.6.

Model assessments on SST, precipitation and surface ocean currents clearly show some simulation improvements after changing the component models. Cold tongue biases are common among the three CGCMs, but the ECHAM5-NEMO3.6 presents warm SST bias in North Pacific with the opposite sign to other two models. Simulation of total radiation doesn't differ much in tropical Pacific, while the surface stress in the ECHAM5-NEMO3.6 and ECHAM6-NEMO3.6 contains intensified eastly anomalies but in the MPI-ESM biases appear in meridional direction. It implies that the momentum anomalies are accountable for cold SST bias in tropical Pacific for coupled models using NEMO, but biases in oceanic dynamics and radiation are accountable for the same cold bias in the MPI-ESM. Further investigation is undertaken on circulations deep into the ocean and atmosphere to get a full picture of bias genesis during the course of coupling.

#### 4.7 Model variability of ENSO and SAM

In the coupled ocean-atmosphere system, global climate variability has been driven by the El Niño-Southern Oscillation (ENSO), the southern annular mode (SAM, also called the Antarctic Oscillation) and the Indian Ocean dipole (IOD) (Philander, 1990; Wallace and Thompson, 2002; Saji et al., 1999). It is therefore necessary to examine the model variability by applying spectra analysis on relevant indices. The CGCM simulations of the three indices are generally consistent with the theoretical red noise (Markov) spectrum (figure omitted). The Niño3.4 index is defined as the SST anomalies averaged over the NINO34 region (5°N-5°S, 170°W -120°W). It shows high variance in 2-7 years' period that documents the ENSO peaks in the HadISST reanalysis (Fig. 8a). All of the CGCMs reproduce similar variations of the Niño3.4

power spectra. The ECHAM6-NEMO3.6 presents weak variabilities at the interannual and interdecadal scale, whose periodic peaks are about one year less than the reanalysis counterpart. The ECHAM5-NEMO3.6 shows a better spectral distribution that best coincides with the reanalysis at the interannual scale. However, it still suffers a weak variability at the interdecadal scale and the periodic peak is even half a year less than that of the ECHAM6-NEMO3.6. The MPI-ESM instead takes on an intensified interannual variability, which stays strong at the interdecadal scale. The Southern Oscillation Index (SOI) is calculated based on the differences in sea level pressure anomalies between Tahiti and Darwin in Australia. In comparison to the Niño3.4 spectra, the SOI exhibits similar peaks at the interannual scale in the ERA-20c reanalysis (Fig. 8b). Nevertheless, all the CGCMs reproduce weak variabilities at the interannual and interdecadal scales. The ECHAM6-NEMO3.6 presents the best simulation with a significant increase in variance around 4 years' period, while the ECHAM5-NEMO3.6 shows the weakest variability at the interannual scale. It implies that the AGCM replacement has an opposite effect on the Niño3.4 and SOI variabilities. The MPI-ESM also show reduced variance from the annual scale and above, quite the opposite to that in the Niño3.4 case. Since the model biases of ENSO variability may be attributed to thermocline feedback and zonal wind variations (Borlace et al., 2013), the reversed changes in the variabilities of Niño3.4 and SOI can be caused by the related oceanic and atmospheric processes. The SAM index is calculated following Gong and Wang (1999) by the differences of normalized monthly zonal mean sea level pressure at 40°S and 65°S. Variations of SAM tend to be more flattened than those of SOI in the ERA-20c reanalysis (Fig. 8c), with prominent fluctuations from biannual to interannual scales. Compared with the reanalysis counterpart, all CGCMs show more power at interannual time scales that represents a robust modulation of the SAM, which is possibly attributed to the semi-annual oscillation (SAO) (Hurrell and van Loon, 1994) and circulation anomalies over Antarctica (Thompson and Solomon, 2002). The ECHAM6-NEMO3.6 presents stronger decadal variability than that of the reanalysis data, while the ECHAM5-NEMO3.6 exhibits weaker low-frequency variability. Since the high variance in low-frequency band represents the upward trend of SAM index at decadal scale (Raphael and Holland, 2006), updating the AGCM can result in a drastic change of long-term climate variability in southern hemisphere. In contrast, the MPI-ESM shows the SAM variability very close to the reanalysis counterpart from 1 year and above, indicating that the OGCM feedback to the atmosphere can lead to a better representation of the inter-decadal variability.

## 5. Circulation Patterns Relevant to SST Biases

The importance of Meridional Overturning Circulation (MOC) to SST in coupled models has been proved in previous studies (Wang et al., 2014, Liu et al., 2016). Thus, the comparison of North Pacific MOC (NPMOC) (Fig. 9) between the CGCMs with the same atmospheric or oceanic component model can help to explain bias characteristics in relation to SST deviations. Model bias pattern can be influenced by the choice of ocean reanalysis dataset. Uncertainties in the ocean reanalyses (ORAs) include the estimation of sea ice thickness, interannual variability of salinity, surface heat flux and mixed layer depth (Balmaseda et al., 2015; Toyoda et al., 2017). There are substantive discrepancies in temperature, salinity and density in the deep

ocean and the Southern Ocean, where observations are sparse especially before the year 2000. Ocean heat content at deep levels varies largely among the ORAs, with majority of spread originating in the Southern Hemisphere due to lack of observation for data assimilation (Palmer et al., 2017). Discrepancies among the ORAs are evident in the strength and structures of the AMOC, with distinctive differences in the depth of equatorward return flow and the depth of the maximum AMOC in northern high latitudes (Karspeck et al., 2017). However, the uncertainties of climatology above bathypelagic zone are less than those of deeper levels. Since the SODA reanalysis data has been employed in many ocean studies, including the North Pacific MOC that is mostly within the thermocline layer, it is reliable to use the SODA reanalysis data to analyse model biases related to the MOC.

### 5.1 Meridional Overturning Circulation

The ECHAM5-NEMO3.6 possesses two prominent improvements of the SST simulation, one in the Pacific cold tongue region, and the other in North Atlantic. It also exhibits warm SST biases in the North Pacific that is the opposite to most CMIP CGCMs. These peculiar characteristics require further investigation on the meridional overturning circulation (MOC) in North Pacific and North Atlantic. For the MOC in North Pacific, the ECHAM5-NEMO3.6 and ECHAM6-NEMO3.6 possess similar bias patterns overall, with intensified tropical cell and deep tropical cell that are mentioned in Liu et al. (2011) (Fig. 9a, b). Tropical cell enhancement is more significant in the ECHAM5-NEMO3.6, so that the upwelling in the tropics and subsequent heat transport to mid-latitudes are more than those in the ECHAM6-NEMO3.6. Bias in the MPI-ESM experiment manifests itself to a larger degree (Fig. 9c). The piControl experiment result (available in <http://esgf-node.llnl.gov/>) is attached (Fig. 9d) to demonstrate that the prominent biases in MPI-ESM are not caused by increasing coupling frequency from one day to 4 hours. The piControl run data used in this study has the same time span as that of the reanalysis data from the year 1981 to 2010, which should be able to represent the model abilities in reproducing the climatology of the same period. More contour lines appear in NPMOC bias distribution in piControl than the MPI-ESM setting in this paper (Fig. 9c, d), suggesting obvious improvements after decreasing coupling interval. This is consistent with previous studies (Bernie et al., 2008; Ge et al., 2017). Enhancement of tropical cell in MPI-ESM is unremarkable, compared with that of ECHAM5-NEMO3.6, but deep tropical cell is significantly intensified that tropical upwelling is forced to become too strong. This may explain why tropical SST bias of MPI-ESM is more than 2°C in boreal summer, without significant radiation errors and easterly anomalies in surface stress. Excessive upwelling in tropical Pacific, induced by intensified deep tropical cell of NPMOC, cools down the local SST. Unlike the ECHAM5-NEMO3.6 case, where the SST cooling in tropical Pacific is driven by intensified tropical cell of NPMOC that transports more heat to mid-latitudes, the MPI-ESM simulation of tropical and subtropical cells does not differ much from SODA reanalysis data. The resulting poleward heat transport carried by NPMOC is not increased in the MPI-ESM experiment. Therefore, the SST bias in North Pacific remains negative (Fig. 3e) under the impact of Northern Hemisphere annular mode (NAM) and wind-evaporation-SST (WES) feedback, when cold SST biases appear in tropical and extratropical North Atlantic (Zhang & Zhao, 2015).

The cold SST bias in North Pacific and North Atlantic for ECHAM6-NEMO3.6 can also be explained with the same reason. But for the ECHAM5-NEMO3.6 case, the poleward heat transport has been enhanced by intensified tropical cell to bring up the SST in mid-latitudes (Fig. 3c).

Since the upper cell of Atlantic meridional overturning circulation (AMOC) plays a significant role in delaying warming signals from anthropogenic greenhouse gases and responding to climate change (Marshall et al., 2014; Buckley and Marshall, 2016), model bias analysis is still focused on the upper ocean levels. The overall magnitude of AMOC bias is less than that of NPMOC with significantly reduced biases near the sea surface (Fig. 10), which is consistent with those of surface currents among the three CGCMs. The ECHAM6-NEMO3.6 shows exiguous bias near the ocean surface, but presents strong biases in the mesopelagic zone of subtropical areas, bringing more heat to higher latitudes (Fig. 10a). Likewise, the ECHAM5-NEMO3.6 exhibits strong circulation biases rotating clockwise in the thermocline that intensifies poleward heat transport (Fig. 10b). With similar bias patterns of the AMOC, the ECHAM5-NEMO3.6 and the ECHAM6-NEMO3.6 have opposite SST biases in North Atlantic (Figs. 3a and 3c), which implies that the air-sea feedback including WES feedback and NAM as suggested by Zhang & Zhao (2015) takes the responsibility. The MPI-ESM experiment shows negative biases in tropical Atlantic from the sea surface to the bathypelagic zone, indicating that the overturning circulation has been restrained. There is a narrow positive bias in the subtropical Atlantic, but its strength has been limited by the negative biases nearby. One consequence of the weak AMOC is the decrease of SST in North Atlantic due to less heat supply from the tropics (Fig. 3e). The overturning circulation is enhanced in the middle latitudes with one centre located north of 35° N and another centre around 55° N at the depth of 1200m. It still promotes the poleward heat transport and results in warm SST biases in subpolar region (Fig. 3e). The AMOC biases in the MPI-ESM piControl experiment are similar as those in the MPI-ESM experiment, with more negative biases in tropical Atlantic. Comparing the AMOC biases between the MPI-ESM and the ECHAM5-NEMO3.6, it can be seen that the SST cold biases in North Atlantic are partially attributed to decreased MOC in the thermocline of tropical and extra-tropical oceans. However, the air-sea interaction also takes account of the SST variations in consideration of the SST differences between the ECHAM5-NEMO3.6 and the ECHAM6-NEMO3.6. Zhang & Zhao (2015) suggested that the cold SST bias in Atlantic caused the same cold bias in North Pacific through different mechanisms originating in tropical and extra-tropical Atlantic. Because the differences of NPMOC are bigger than those of AMOC between these two newly developed CGCMs, it suggests an inverse cause-and-effect relationship between the cold SST biases in North Pacific and North Atlantic where the former takes the lead.

## 5.2 Vertical Structure of Atmospheric Circulation

On behalf of the atmosphere motion that accounts for cold tongue bias and strong easterly bias in surface currents, vertical circulation bias zonally averaged over South Equatorial Current (SEC) against the ERA-Interim reanalysis is given in Fig. 9. Consistent with surface wind stress bias (Fig. 5), there are easterly anomalies in the lower atmosphere for ECHAM5-NEMO3.6 and ECHAM6-NEMO3.6 (Fig. 11a, b). With



anomalous upward (downward) motion over the western (eastern) Pacific, the easterly anomalies are maintained by the anomalous Walker circulation across tropical Pacific. The easterly anomalies in lower atmosphere decrease significantly in MPI-ESM experiment (Fig. 11c), without strong downward motion in the eastern Pacific. Inter-model differences in the easterly biases at low levels are consistent with the amplitude of NPMOC tropical cell, because the stronger surface winds collaborating with Coriolis effect push more sea water to move west and poleward.

It can be seen from Figure 11 that cold temperature biases in the MPI-ESM transform into warm temperature biases in the ECHAM5-NEMO3.6 accompanied by increasing easterly anomalies in lower atmosphere, which suggests warmer temperature bias is correlated with stronger easterly anomalies. The pattern correlation between temperature biases and easterly anomalies in the three CGCM experiments has been calculated respectively. Because the easterly biases are most remarkable below 800hPa, the levels taken into consideration are between 800hPa and 1000hPa. The pattern correlation coefficients between temperature biases and zonal wind biases are 0.315 for the ECHAM6-NEMO3.6 experiment, 0.586 for the ECHAM5-NEMO3.6 experiment, and 0.411 for the MPI-ESM experiment. All of them have passed the 99% Student-t significance test. The ECHAM5-NEMO3.6 has the most significant easterly biases and holds the largest correlation coefficient, much bigger than that of the other two CGCMs, which confirms the correlation between the warmer temperature and stronger easterly anomalies. Recalling precipitation biases over the SEC area (Fig. 4a, c, e), where ECHAM5-NEMO3.6 shows the driest situation and MPI-ESM presents the wettest circumstance, the latent heat absorption in the course of evaporation turns out to be responsible for the temperature bias. More precipitation requires higher specific humidity that favours cloud formation, in which more latent heat of vaporization is taken up when water vapor mixing ration is increased. Since cumulus convection modulates changes in temperature, specific humidity and atmospheric circulation, it is most likely to be the predominant factor that shapes the inter-model differences.

The analysis on oceanic and atmospheric circulation has made it clear that the SST bias is consistent with meridional overturning circulation in North Pacific, driven by surface wind stress anomalies that are maintained by anomalous Walker circulation over the tropical Pacific. Cumulus convection process is found to be a major contributor to inter-model differences. To better understand the impact by changing each component model, it is necessary to quantitatively analyse their contributions to simulation differences.

## 6. Contribution of Each Component Model

A total of 11 variables that are sensitive to air-sea coupling are selected to calculate pattern correlation with SST simulation differences in boreal summer. Correlations between model differences in SST and each variable with ECHAM6-NEMO3.6 minus ECHAM5-NEMO3.6, denoted as “AGCM” column in Table 1, indicate the effects of changing atmospheric component model at the global scale. Similarly, correlation results with ECHAM6-NEMO3.6 minus MPI-ESM to represent the contribution of changing oceanic component model are named “OGCM” in Table 1. Because the analysis in previous sections is mainly focused on the Pacific, the pattern correlation over the area is also provided for comparison (Table 2).

## 6.1 Effects of AGCM Replacement

For the AGCM case, net longwave radiation is ranked top in the global correlation, followed by surface evaporation (latent heat flux) and net shortwave radiation (Tab. 1). When the study area is narrowed down to the Pacific, the top 3 ranking variables are the same, with the net shortwave radiation takes the 1<sup>st</sup> place (Tab. 2). In both cases, the net shortwave and net longwave fluxes are ranked in top 3, which suggests that radiation budget is closely associated with SST inter-model differences. The pattern correlations of the 3 variables are higher within the Pacific than that at the global scale, indicating an increased robustness of the variable estimates. Changing the atmospheric component model from ECHAM-6.3 to ECHAM-5.4 affects cumulus convection processes, including temperature, specific humidity and winds, which further alter the ground radiation fluxes through cloud radiation feedback. Noting that latent heat flux has a correlation value close to that of net longwave radiation, it would suggest that cumulus convection changes sea surface winds and then surface evaporation, leading to differences in latent heat flux. Consequently, the SST is altered in tropical oceans through wind-evaporation-SST (WES) feedback. The negative sign of correlations for longwave and latent heat flux (Tab. 1) suggests a contrary trend with SST variations. It is easy to understand because more latent heat and longwave dissipation cools down the surface sea water. Similarly, positive correlation for shortwave flux is due to more solar irradiance that brings up the SST.

To determine the physical processes responsible for simulation differences, changes in SST, surface winds, radiation budget and vertical circulation are plotted, respectively. It can be seen from Fig. 12d that negative deviations almost occupy the northern hemisphere, including tropical Pacific and Indian Ocean, North Pacific and North Atlantic. First, the mechanisms behind SST deviations in tropical and subtropical oceans are discussed here. Deviations of 10m wind exhibit northerly anomalies east of dateline in southern tropical Pacific (Fig. 12b), where easterly winds prevail for summer climatology (Fig. 10a), hence the evaporation and latent heat absorption over the sea surface are enhanced that makes SST deviation colder than 1°C (Fig. 12d). There are eastward oriented wind anomalies over subtropical Indian Ocean, subtropical Atlantic and some parts of subtropical Pacific (Fig. 12b) in the opposite direction against climatology (Fig. 12a). Their superposition results in a decreased wind speed that reduces surface evaporation and latent heat flux, which finally keeps the SST warm for those sea waters. The SST deviations thus form a positive feedback (WES feedback) with changes of surface wind and evaporation, in accordance with signs of correlation in Table 1. Since the latent heat and surface wind differences are caused by replacing the AGCM, it is advisable to compare deviations in vertical circulation that may shed some light on corresponding physical processes. An anti-clockwise Walker circulation accompanied by negative temperature deviations occupies central and eastern Pacific over the SEC area (Fig. 13a). Low-level easterly anomalies are consistent with surface wind distribution (Fig. 12b) because northerly flows in southern tropical Pacific that contributes to more evaporation are not considered in meridional average. It can be assumed that changing the AGCM alters radiation budget, and tropospheric temperature becomes colder so that more air currents cool down and sink by insufficient heat. Downward motion over the SEC area connects easterly anomalies in the middle troposphere, which forms a complete Walker cell that further enhances low-level westerly anomalies.

To verify this assumption, differences in radiation budget with respect to net shortwave and longwave fluxes, latent heat and sensible heat are drawn in Fig. 14. It turns out that deviations in shortwave flux and latent heat are more significant than those in longwave and sensible heat fluxes. Differences in shortwave radiation can be attributed to cloud radiation feedback, which is induced by changes in cumulus convection after replacing the AGCM. Whereas the deviations in latent heat flux are derived from surface wind anomalies. More latent heat release appears over the areas where the surface winds are intensified by local anomalies (Fig. 12a, b). The deviation patterns are congruent with the correlation signs (Tab. 1 and 2). More solar irradiance corresponds to higher SST of subtropical oceans in southern hemisphere, while less shortwave flux matches cold SST in central Pacific. Under the same cloud radiative forcing, net longwave flux shows approximately inverse variations. Differences in surface winds and heat budget also affect the latent heat flux. Therefore, it can be confirmed that the AGCM replacement first alters cumulus convection that modulates temperature, specific humidity and atmospheric circulation, which in turn accommodates cloud radiation feedback to a consistent change and affects the radiation budget.

## 6.2 Impacts of OGCM Replacement

For simulation differences after changing the OGCM, latent heat flux holds the biggest share in pattern correlation with corresponding SST differences (Tab. 1), much bigger than that of net longwave flux at the second place and sensible heat flux at the third place. When the study area is confined within the Pacific, the three leading variables are the same but the sensible heat flux is ranked top (Tab. 2). Compared with the rankings in the AGCM case, it can be seen that the OGCM influences the simulation with different physical mechanisms than those by replacing AGCM. The effect of cloud radiation feedback is not prominent, while the physical processes associated with latent heat and sensible heat, including heat conduction, evaporation and convection, take a bigger share in the SST inter-model differences. SST differences can also lead to changes in longwave flux. The negative sign of correlations for latent heat, sensible heat and longwave flux (Tab. 1 and 2) suggests a reverse trend to SST variations.

Contrary to the AGCM case, low-level wind deviations in tropical Pacific (Fig. 12c) are in the same direction as the background climatology (Fig. 12a), which should increase surface evaporation through stronger near surface winds and then results in colder SST through WES feedback. Nevertheless, warmer SST appears in most parts of tropical and subtropical oceans, laying waste to the assumption trying to explain SST deviation in terms of WES mechanism. Since the top 3 variables that are most relevant to SST deviations (Tab. 1) suggest changes in the surface heat budget, involving the momentum and temperature exchanges between ocean and atmosphere, it implies that the joint effects of atmosphere and ocean models lead to the model deviation patterns (Fig. 15). Simulation of ocean dynamics is different after replacing the OGCM. With changes in ocean advection, the SST and surface currents are altered which modulates the surface evaporation, convection and heat conduction. Subsequently the latent heat and sensible heat fluxes vary over the sea surface. The thermal and moisture perturbations from the ocean are passed to the atmosphere during coupling processes (Fig. 15c, d). Variations of low-level atmospheric circulation and humidity take effects

on cloud formation and cloud liquid water path that changes precipitation and cloud radiative forcing. The net shortwave and longwave radiations are influenced and make a difference to the atmospheric circulation (Fig. 12c) and the heat budget over the sea surface (Fig. 15 a, b). Then, the perturbation signal is transferred back to the ocean that changes the SST and surface currents. This air-sea feedback finally reaches a quasi-equilibrium with marked SST warming over vast maritime spaces over the globe. The associated physical processes represented by each radiation term in Figure 15 are in accordance with their signs of pattern correlation (Tab. 1). It seems to suggest that surface evaporation plays a predominant role in the SST differences, because the latent heat flux holds the biggest correlation coefficient and its deviation pattern is the most obvious among the four radiation terms. However, it is more likely a manifestation of the large-scale ocean dynamical effect on the inter-model differences as suggest by Ying and Huang (2016).

Atmospheric circulation accommodates itself to changes of radiation budget. An anomalous Walker circulation cell appears over the SEC area (Fig. 13b), with updraft over Philippine sea where positive SST deviations indicate a net radiation surplus. The upward flow rises to the upper-level troposphere and condensation becomes evident due to temperature drops. It then diverts toward east and orients downward over eastern Pacific, which enhances surface westerly anomalies in the lower atmosphere. This circulation pattern is consistent with anomalous easterlies over tropical Pacific (Fig. 12c). Warm temperature deviations (Fig.13b) can be viewed as the manifestation of surface radiation surplus and latent heat release of condensation. So far, the analysis has made it clear that changing the OGCM affects the SST simulation through large-scale air-sea feedback that mainly involves surface evaporation, heat conduction, atmospheric convection and cloud masking of incoming and outgoing radiative fluxes. Under the impact of net radiation surplus over western Pacific, the temperature rises with upward motion which forms a warm advection heading east and descending over eastern Pacific. This anomalous Walker cell drives low-level winds towards the west, leading to westerly anomalies over vast areas of tropical Pacific.

## 7. Summary and Discussion

In this study, two new CGCMs have been developed based on the coupling structure of MPI-ESM, namely ECHAM5-NEMO3.6 and ECHAM6-NEMO3.6. The new CGCMs show some improvements in the simulation of SST, precipitation and ocean currents compared with MPI-ESM. The ECHAM5-NEMO3.6 presents the best SST simulation in summer with the minimum cold tongue bias in tropical Pacific and no remarkable bias in North Atlantic, while ECHAM6-NEMO3.6 reproduces the best winter climatology with most of the biases less than 1 °C. For precipitation simulation, ECHAM6-NEMO3.6 presents the highest pattern correlation with reanalysis counterpart and substantially ameliorates double ITCZ problem over SPCZ in boreal summer. Biases in surface currents and meridional overturning circulation are also considerably reduced in both ECHAM5-NEMO3.6 and ECHAM6-NEMO3.6. Wind stress bias patterns are alike in most areas among the three CGCMs, but MPI-ESM shows only poleward anomalies over tropical Pacific without strong easterly biases that are common in most existing coupled models. All three CGCMs can generally reproduce the ENSO and SAM variabilities. Model biases are more evident at the interannual

and interdecadal scales, suggesting a prominent effect of circulation and oscillation anomalies in both atmosphere and ocean models on the air-sea feedback (Yu and Kim, 2011; Farneti, et al. 2014). As suggested by Cai et al. (2011), the evolution of these climate drivers is affected by their interactions, which requires further investigation. The ECHAM5-NEMO3.6 has much larger biases in total radiation than MPI-ESM, whereas the latter presents the maximum deviation in SST simulation. Besides, the ECHAM5-NEMO3.6 shows warm SST bias in North Pacific, with the opposite sign to most CGCM biases at present time. These facts constitute evidence that suggests model errors of each CGCM are caused by different physical mechanisms.

Meridional overturning circulation in North Pacific and vertical structure of atmospheric circulation are analysed for a comprehensive understanding of bias genesis in each CGCM. Overestimation of tropical cell in the NPMOC transfers more heat to mid-latitudes and results in warm SST bias in North Pacific of the ECHAM5-NEMO3.6. Excessively strong deep tropical cell of the MPI-ESM intensifies the tropical upwelling that leads to the largest cold SST bias among all three CGCMs. The analysis on atmospheric vertical circulation over the SEC area has confirmed that momentum field plays a major role in the SST biases of ECHAM5-NEMO3.6 and ECHAM6-NEMO3.6, while oceanic processes and radiation budget are responsible for the same cold SST biases in tropical Pacific. Since the inter-model differences are caused by changing component models, 12 surface variables that are sensitive to coupling processes are chosen to calculate pattern correlation with the SST differences between CGCMs with the same atmospheric or oceanic component model. The top 3 variables ranked in the case of changing the AGCM are net longwave radiation, surface evaporation (latent heat flux) and net shortwave radiation. Differences in SST, radiation budget and atmospheric circulation are analysed, and it is confirmed that the AGCM replacement first militates in the alteration of cumulus convection including temperature, specific humidity and atmospheric circulation, which in turn changes SST through WES feedback and affects the radiation budget through cloud radiation feedback. For the OGCM replacement, latent heat flux (evaporation) holds the maximum pattern correlation with SST variation, much larger than that of net longwave flux at the 2<sup>nd</sup> place and sensible heat at the 3<sup>rd</sup> place. Through analysis on circulation and radiation terms, it has been clear that the ocean dynamical effect plays a predominant role in the SST differences after changing the OGCM. The ocean advection initiates the perturbations of SST and surface evaporation that modulate the atmospheric humidity and low-level circulation. Consequently, the cloud masking effect on radiative fluxes are altered which influences the atmospheric circulation and surface heat budget. The resulting equilibrium of the air-sea feedback manifest itself as the inter-model differences in the related meteorological fields. In the OGCM case, the latent heat flux holding the largest correlation coefficient at the global scale is consistent with Cao et al. (2015), which points out that amplitude and meridional variability of latent heat flux over Pacific are the most diverse in CMIP5 models.

It is noteworthy that the SST deviations by changing the OGCM and AGCM are nearly the opposite (Fig. 12d, e), suggesting that reverse transformation of model bias can be realized through mechanisms unveiled in this paper. Although the top 3 coupling variables that are most relevant to SST deviations after

changing the AGCM or OGCM are radiation terms, the physical mechanisms behind the opposite SST variations are different. AGCM replacement affects cumulus convection that eventually changes momentum field and radiation budget over the sea surface, while OGCM replacement alters sea surface evaporation that results in latent heat variations and consequently leads to readjustment of radiation budget. It implies that one can pursue simulation improvement of the cold tongue from two aspects: 1. Improve the cumulus convection scheme in the AGCM, 2. Ameliorate errors of surface radiation budget in the OGCM, especially for latent heat. With better cumulus convection scheme, easterly anomalies over the tropical sea surface can be reduced significantly, which will cut down the latent heat absorption of more evaporation induced by stronger surface winds, and consequently break down the WES feedback that amplifies the cold SST bias. For the OGCM part, better treatment of surface evaporation and radiation budget can reduce the cold SST bias in relation to net radiation deficiency.

For strong easterly bias over eastern Pacific in ECHAM6-NEMO3.6, also common in the most CGCMs, the MPI-ESM instead shows poleward bias, which suggests that OGCM replacement can also diminish this bias through coupling processes. It is easy to see that an anomalous Walker circulation with counter-clockwise rotation will appear over the tropical Pacific, if net surface radiation is warmer in the east and colder in the west. With proper OGCM configuration, sea surface evaporation can initiate this radiation anomaly, and the resulting Walker circulation will decrease the easterly wind speed. Although the analysis in section 5 and 6 only focuses on the physical mechanisms behind simulation differences in summer, the atmospheric and oceanic circulations have a similar impact on radiation budget and surface heat transport in other seasons. It implies that the effects by replacing each component model may be the same, and the conclusions drawn above are also valid for SST biases in winter. The mechanisms illustrated in this study provides a new vision of model bias origin, which is heuristic for model improvement researches and practices.

## Code and data availability

The model source code is available from the authors upon request. Restrictions on source code license will be imposed unless for academic and non-commercial use. The experimental data can be found in <https://doi.org/10.5281/zenodo.1306338> (Gui et al., 2018).

The reanalysis data used in this study can be downloaded from the following websites:

1. The Hadley Centre SST data is downloaded from <https://www.metoffice.gov.uk/hadobs/hadisst/>.
2. The GPCP precipitation data is downloaded from <https://precip.gsfc.nasa.gov/>.
3. The ERA-Interim and ERA-20c monthly reanalysis data is available at <http://apps.ecmwf.int/datasets/>.
4. The SCOW wind stress data is available at <http://cioss.coas.oregonstate.edu/scow>.
5. Surface radiation reanalysis data of CERES EBAF-Surface Ed4.0 is downloaded from <https://ceres.larc.nasa.gov>.
6. The SODA3.3 ocean current reanalysis data is downloaded from: <http://www.atmos.umd.edu>.

The MPI-ESM piControl experiment data is available at <http://esgf-node.llnl.gov>.

**Acknowledgements**

This work was supported by the National Key Research and Development Program of China (2016YFA0601600), and the National Natural Science Foundation of China (U1502233 and 41565002). It is also co-supported by Yunnan University's Research Innovation Fund for Graduate Students (YDY17019).



## References

- Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R. R., Xie, P., Janowiak, J. E., ... Nelkin, E.: The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present), *J. Hydrometeorol.*, 4, 1147–1167, 2003.
- Baehr, J., Fröhlich, K., Botzet, M., Domeisen, D. I. V., Kornblueh, L., Notz, D., Piontek, R., Pohlmann, H., Tietsche, S., and Müller, W. A.: The prediction of surface temperature in the new seasonal prediction system based on the MPI-ESM coupled climate model, *Clim. Dynam.*, 44, 2723, doi:10.1007/s00382-014-2399-7, 2015.
- Balmaseda, M. A., Hernandez, F., Storto, A., Palmer, M. D., Alves, O., Shi, L., Smith, G. C., Toyoda, T., Valdivieso, M., Barnier, B., Behringer, D., Boyer, T., Chang, Y.S., Chepurin, G. A., Ferry, N., Forget, G., Fujii, Y., Good, S., Guinehut, S., Haines, K., Ishikawa, Y., Keeley, S., Kohl, A., Lee, T., Martin, M. J., Masina, S., Masuda, S., Meyssignac, B., Mogensen, K., Parent, L., Peterson, K. A., Tang, Y. M., Yin, Y., Vernieres, G., Wang, X., Waters, J., Wedd, R., Wang, O., Xue, Y., Chevallier, M., Lemieux, J. F., Dupont, F., Kuragano, T., Kamachi, M., Awaji, T., Caltabiano, A., WilmerBecker, K. and Gaillard, F.: The Ocean Reanalysis Intercomparison project (ORAIP), *J. OPER. OCEANOGR.*, 8(sup1): s80-s97, doi: 10.1080/1755876X.2015.1022329, 2015.
- Bernie, D., Guilyardi, E., Madec, G., Slingo, J., Woolnough, S., and Cole, J.: Impact of resolving the diurnal cycle in an ocean–atmosphere GCM. Part 2: a diurnally coupled CGCM, *Clim. Dynam.*, 31, 909–925, 2008.
- Bond, N. A. and Cronin, M. F.: Regional weather patterns during periods of anomalous air–sea interactions at the Kuroshio Extension Observatory (KEO), *J. Climate*, 21, 1680–1697, 2008.
- Borlace, S., Cai, W., Santoso, A.: Multidecadal ENSO amplitude variability in a 1000-yr simulation of a coupled global climate model: implications for observed ENSO variability, *J. Climate*, 26, 9399–9407, 2013.
- Bouillon, S., Fichet, T., Legat, V., and Madec, G.: The elastic–viscous–plastic method revisited, *Ocean Model.*, 71, 2–12, 2013.
- Brovkin, V., Boysen, L., Raddatz, T., Gayler, V., Loew, A., and Claussen, M.: Evaluation of vegetation cover and land-surface albedo in MPI-ESM CMIP5 simulations, *J. Adv. Model. Earth Sy.*, 5, 48–57, 2013.
- Brovkin, V., Raddatz, T., Reick, C., Claussen, M., and Gayler, V.: Global biogeophysical interactions between forest and climate, *Geophys. Res. Lett.*, 36, L07405, doi:10.1029/2009GL037543, 2009.
- Buckley, M. W., and Marshall, J.: Observations, inferences, and mechanisms of the Atlantic Meridional Overturning Circulation: A review, *Rev. Geophys.*, 54, doi: 10.1002/2015RG000493, 2016.
- Burls, N. J., Muir, L., Vincent, E. M., and Fedorov, A.: Extra-tropical origin of equatorial Pacific cold bias in climate models with links to cloud albedo, *Clim. Dynam.*, 49, 2093, doi:10.1007/s00382-016-3435-6, 2017.
- Cai, W., Sullivan, A., and Cowan, T.: Interactions of ENSO, the IOD, and the SAM in CMIP3 Models, *J. Climate*, 24, 1688–1704, doi:10.1175/2010JCLI3744.1, 2011.

- Cao, N., Ren, B. H., and Zheng, J. Q.: Evaluation of CMIP5 climate models in simulating 1979–2005 oceanic latent heat flux over the Pacific, *Adv., Atmos. Sci.*, 32, 1603–1616, 2015.
- Carton, J. A., and Giese, B. S.: A Reanalysis of Ocean Climate Using Simple Ocean Data Assimilation (SODA), *Mon. Weather. Rev.*, 136, 2999–3017, 2008.
- Chen, S. S., Zhao, W., Donelan, M. A., and Tolman H. L.: Directional Wind–Wave Coupling in Fully Coupled Atmosphere–Wave–Ocean Models: Results from CBLAST-Hurricane, *J. Atmos. Sci.*, 70, 3198–3215, doi:10.1175/JAS-D-12-0157.1, 2013.
- Craig A., Valcke S., Coquart L.: Development and performance of a new version of the OASIS coupler, OASIS3-MCT\_3.0, *Geosci., Model Dev.*, 10, 3297–3308, doi:10.5194/gmd-10-3297-2017, 2017.
- Dewitte, B., Thual, S., Yeh, S., An, S., Moon, B., and Giese, B.S.: Low-Frequency Variability of Temperature in the Vicinity of the Equatorial Pacific Thermocline in SODA: Role of Equatorial Wave Dynamics and ENSO Asymmetry, *J. Climate*, 22, 5783–5795, doi:10.1175/2009JCLI2764.1, 2009.
- Drenkard, E.J. and Karnauskas, K.B.: Strengthening of the Pacific Equatorial Undercurrent in the SODA Reanalysis: Mechanisms, Ocean Dynamics, and Implications, *J. Climate*, 27, 2405–2416, doi:10.1175/JCLI-D-13-00359.1, 2014.
- Dunlap, R., Vertenstein, M., Valcke, S., and Craig, T.: Second Workshop on Coupling Technologies for Earth System Models. *B. Am. Meteorol. Soc.*, 95, ES34–ES38, doi:10.1175/BAMS-D-13-00122.1, 2014.
- Fan, Y. L., and Hwang, P. Kinetic energy flux budget across air-sea interface, *Ocean Model.*, 120, 27–40, doi:10.1016/j.ocemod.2017.10.010, 2017.
- Farneti, R., Molteni, F., Kucharski, F.: Pacific interdecadal variability driven by tropical-extratropical interactions, *Clim. Dynam.*, 42: 3337–3355, 2014.
- Fichefet, T. and Morales Maqueda, M. A. Sensitivity of a global sea ice model to the treatment of ice thermodynamics and dynamics, *J. Geophys. Res.*, 102, 12609–12646, 1997.
- Ge, X., Wang, W., Kumar, A. and Zhang, Y.: Importance of the Vertical Resolution in Simulating SST Diurnal and Intraseasonal Variability in an Oceanic General Circulation Model, *J. Climate*, 30, 3963–3978, doi:10.1175/JCLI-D-16-0689.1, 2017.
- Gong, D. and Wang, S.: Definition of Antarctic oscillation index, *Geophys. Res. Lett.*, 26, 459–462, doi:10.1029/1999GL900003, 1999.
- Gualdi, S., Navarra, A., Guilyardi, E., and Delecluse, P.: Assessment of the tropical Indo-Pacific climate in the SINTEX CGCM, *Ann. Geophys-italy.*, 46,1, doi: 10.4401/ag-3385, 2003.
- Gui S., Yang R.W., and Cao J.: Gui et al. Experimental data, <https://doi.org/10.5281/zenodo.1306338>, 2018.
- Ham, Y. and Kug, J.: Improvement of ENSO Simulation Based on Intermodel Diversity, *J. Climate*, 28, 998–1015, doi:10.1175/JCLI-D-14-00376.1, 2015.
- Huang, P., Wang, P.F., Hu, K.M., Huang, G., Zhang, Z.H., Liu, Y., and Yan, B.L.: An Introduction to the Integrated Climate Model of the Center for Monsoon System Research and Its Simulated Influence of El Nino on East Asian–Western North Pacific Climate, *Adv. Atmos. Sci.*, 31, 1136–1146, 2014.

- Johnson, G. C., Schmidtko, S., and Lyman, J. M.: Relative contributions of temperature and salinity to seasonal mixed layer density changes and horizontal density gradients, *J. Geophys. Res.*, 117, C04015, doi:10.1029/2011JC007651, 2012.
- Jungclaus, J. H., Keenlyside, N., Botzet, M., Haak, H., Luo, J. J., Latif, M., Marotzke, J., Mikolajewicz, U., and Roeckner, E.: Ocean circulation and tropical variability in the coupled model ECHAM5/MPI-OM, *J. Climate*, 19, 3952–3972, 2006.
- Jungclaus, J. H., Fischer, N., Haak, H., Lohmann, K., Marotzke, J., Matei, D., Mikolajewicz, U., and Von Storch, J. S.: Characteristics of the ocean simulations in MPIOM, the ocean component of the MPI Earth System Model. *J. Adv. Model. Earth Sy.*, 5, 422–446, doi:10.1002/jame.20023, 2013.
- Kanzow, T., Cunningham, S. A., Johns, W. E., Hirschi, J. J., Marotzke, J., Baringer, M. O., Meinen, C. S., Chidichimo, M. P., Atkinson, C., Beal, L. M., Bryden, H. L., and Collins, J.: Seasonal variability of the Atlantic meridional overturning circulation at 26.5°N, *J. Climate*, 23, 5678–5698, doi:10.1175/2010JCLI3389.1, 2010.
- Karspeck, A. R., Stammer, D., Köhl, A., Danabasoglu, G., Balmaseda, M., Smith, D. M., Fujii, Y., Zhang, S., Giese, B., Tsujino, H., and Rosati, A.: Comparison of the Atlantic meridional overturning circulation between 1960 and 2007 in six ocean reanalysis products, *Clim. Dyn.*, 49: 957, doi: 10.1007/s00382-015-2787-7, 2017.
- Kato, S., Loeb, N. G., Rose, F. G., Doelling, D. R., Rutan, D. A., Caldwell, T. E., Yu, L., and Weller, R. A.: Surface irradiances consistent with CERES-derived top-of-atmosphere shortwave and longwave irradiances, *J. Climate*, 26, 2719–2740, doi:10.1175/JCLI-D-12-00436.1, 2013.
- Klein, S., and Hartmann, D. L.: The seasonal cycle of low stratiform clouds. *J. Climate*, 6, 1587–1606, 1993.
- Lee, T., Waliser, D. E., Li, J. F., Landerer, F. W., and Gierach, M. M.: Evaluation of CMIP3 and CMIP5 Wind Stress Climatology Using Satellite Measurements and Atmospheric Reanalysis Products, *J. Climate*, 26, 5810–5826, doi:10.1175/JCLI-D-12-00591.1, 2013.
- Li, J. L., Lee, W., Waliser, D. E., Neelin, J. D., Stachnik, J. P., and Lee, T.: Cloud-precipitation-radiation-dynamics interaction in global climate models: A snow and radiation interaction sensitivity experiment, *J. Geophys. Res.*, 119, 3809–3824, doi:10.1002/2013JD021038, 2014.
- Li, G., Xie, S. P.: Tropical biases in CMIP5 multimodel ensemble: the excessive equatorial Pacific cold tongue and double ITCZ problems, *J. Climate*, 27, 1765–1780, doi:10.1175/JCLI-D-13-00337.1, 2014.
- Lin, J. L.: The double-ITCZ problem in IPCC AR4 coupled GCMs: ocean-atmosphere feedback analysis, *J. Climate*, 20, 4497–4525, doi:10.1175/JCLI4272.1, 2007.
- Lin, K., Yang, S., and Chen, S. S.: Reducing TC Position Uncertainty in an Ensemble Data Assimilation and Prediction System: A Case Study of Typhoon Fanapi (2010), *Weather Forecast.*, 33, 561–582, doi:10.1175/WAF-D-17-0152.1, 2018.
- Liu Z., Philander S. G. H., and Pacanowski R. C.: A GCM study of tropical-subtropical upper-ocean water exchange, *J. Phys. Oceanogr.*, 24, 2606–2623, 1994.

- Liu, H., Zhang, Q., Duan, Y., and Hou, Y.: The three-dimensional structure and seasonal variation of the North Pacific meridional overturning-circulation, *Acta Oceanol. Sin.*, 30: 33-42, 2011.
- Liu, H., Zhang, Q., Hou, Y., and Duan, Y.: Interannual variability in the North Pacific meridional overturning circulation. *Chin. J. Oceanol. Limn.*, 31, 665-680, 2013.
- Liu J., Zhou T.J., Wu C., and Guo Z.: Water vapor and cloud radiative feedback processes in the ocean-atmosphere coupled model FGOALS\_g1, *Chinese Journal of Atmospheric Sciences*, 35, 531–546, doi:10.3878/j.issn.1006-9895.2011.03.13, 2011.
- Luo, J., Masson, S., Roeckner, E., Madec, G., and Yamagata, T.: Reducing climatology bias in an ocean-atmosphere CGCM with improved coupling physics, *J. Climate*, 18, 2344–2360, doi:10.1175/JCLI3404.1, 2005.
- Maier-Reimer, E., Mikolajewicz, U., and Hasselmann, K.: Mean circulation of the Hamburg LSG OGCM and its sensitivity to the thermohaline surface forcing, *J. Phys. Oceanogr.*, 23, 731–757, 1993.
- Marshall, J., Armour, K. C., Scott, J. R., Kostov, Y., Hausmann, U., Ferreira, D., Shepherd, T. G., and Bitz, C. M.: The ocean's role in polar climate change: Asymmetric Arctic and Antarctic responses to greenhouse gas and ozone forcing, *Philos. Trans. R. Soc. A*, 372(2019), doi:10.1098/rsta.2013.0040, 2014.
- Marsland, S. J., Haak, H., Jungclaus, J. H., Latif, M., and Röske, F.: The Max Planck Institute global ocean/sea ice model with orthogonal curvilinear coordinates, *Ocean Model.*, 5, 91–127, 2003.
- McCreary J., and Yu Z.: Equatorial dynamics in the 2.5 layer model, *Prog. Oceanogr.*, 29, 61-132, 1992.
- Megann, A. P., Storkey, D., Aksenov, Y., Alderson, S. G., Calvert, D., Graham, T., Hyder, P., Siddorn, J., and Sinha, B.: GO5.0: the joint NERC–Met Office NEMO global ocean model for use in coupled and forced applications, *Geosci. Model Dev.*, 7, 1069-1092, doi: 10.5194/gmd-7-1069-2014, 2014.
- Mobis, B., and Stevens, B.: Factors controlling the position of the intertropical convergence zone on an aquaplanet, *J. Adv. Model. Earth Sy.*, 4, M00A04, doi:10.1029/2012MS000199, 2012.
- Mochizuki, T. & Awaji, T.: Summertime Evolution of Decadal Sea Surface Temperature Anomalies in the Midlatitude North Pacific, *J. Climate*, 21, 1569–1588, doi:10.1175/2007JCLI1853.1, 2008.
- Mogensen, K., Balmaseda, M. A., and Weaver, A.: The NEMOVAR ocean data assimilation system as implemented in the ECMWF ocean analysis for System 4. ECMWF Tech. Memo., 668, 59, <http://old.ecmwf.int/publications/library/ecpublications/pdf/tm/601-700/tm668.pdf>, 2012a.
- Norris, J.R., and Leovy, C.B.: Interannual variability in stratiform cloudiness and sea surface temperature, *J. Climate*, 7, 1915–1925, 1994.
- Norris, J. R., Zhang, Y., and Wallace, J. M.: Role of low clouds in summertime atmosphere–ocean interactions over the North Pacific, *J. Climate*, 11, 2482–2490, 1998.
- Notz, D., Haumann, F. A., Haak, H., Jungclaus, J. H., and Marotzke, J.: Arctic sea-ice evolution as modelled by MPI-ESM, *J. Adv. Model. Earth Sy.*, doi:10.1002/jame.20016, 2013.
- Pham, H. T., Smyth, W. D., Sarkar, S., and Moum, J. N.: Seasonality of Deep Cycle Turbulence in the Eastern Equatorial Pacific. *J. Phys. Oceanogr.*, 47, 2189–2209, doi:10.1175/JPO-D-17-0008.1, 2017.

- Pacanowski, R. C. and Philander, S. G. H.: Parameterization of vertical mixing in numerical models of tropical oceans, *J. Phys. Oceanogr.*, 11, 1443–1451, 1981.
- Palmer, M. D., Roberts, C. D., Balmaseda, M., Chang, Y.-S., Chepurin, G., Ferry, N., Fujii, Y., Good, S. A., Guinehut, S., Haines, K., Hernandez, F., Köhl, A., Lee, T., Martin, M. J., Masina, S., Masuda, S., Peterson, K. A., Storto, A., Toyoda, T., Valdivieso, M., Vernieres, G., Wang, O., and Xue, Y., Ocean heat content variability and change in an ensemble of ocean reanalyses, *Climate Dynamics*, 49: 909–930, doi:10.1007/s00382-015-2801-0, 2017.
- Park, W., Keenlyside, N., Latif, M., Stroh, A., Redler, R., Roeckner, E., and Madec, G.: Tropical Pacific climate and its response to global warming in the Kiel climate model, *J. Climate*, 22, 71–92, 2009.
- Philander, S. G.: *El Niño, La Niña, and the Southern Oscillation*, Academic Press, 289, 1990.
- Raddatz, T., Reick, C., Knorr, W., and Kattge, J.: Will the tropical land biosphere dominate the climate–carbon cycle feedback during the twenty-first century? *Clim. Dynam.*, 29, 565–574, doi:10.1007/s00382-007-0247-8, 2007.
- Raphael, M. and Holland, M. M.: Twentieth Century Simulation of the Southern Hemisphere in Coupled Models. Part I: Large scale Circulation Variability, *Clim. Dynam.*, 26, 217–228, 2006.
- Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *J. Geophys. Res.*, 108, 4407, 2003.
- Redi, M. H.: Oceanic isopycnal mixing by coordinate rotation, *J. Phys. Oceanogr.*, 12, 1154–1158, 1982.
- Reick, C. H., Raddatz, T., Brovkin, V., and Gayler, V.: Representation of natural and anthropogenic land cover change in MPI-ESM, *J. Adv. Model. Earth Sy.*, 5, 459–482, doi: 10.1002/jame.20022, 2013.
- Ren, R., Yang, Y., Cai, M., and Rao, J.: Understanding the systematic air temperature biases in a coupled climate system model through a process-based decomposition method, *Clim. Dynam.*, 45, 1801–1817, doi:10.1007/s00382-014-2435-7, 2015.
- Risien, C. M., and Chelton, D. B.: A global climatology of surface wind and wind stress fields from eight years of QuikSCAT scatterometer data, *J. Phys. Oceanogr.*, 38(11), 2379–2413, 2008.
- Roeckner, E., Dümenil, L., Kirk, E., Lunkeit, F., Ponater, M., Rockel, B., Sausen, R., and Schlese, U.: The Hamburg version of the ECMWF model (ECHAM), *World Meteorol. Tech. Rep.* 13, 1989.
- Roeckner, E., Arpe, K., Bengtsson, L., Christoph, M., Claussen, M., Dumenil, L., Esch, M., Giorgetta, M. A., Schlese, U., Schulzweida, U.: The atmospheric general circulation model ECHAM-4: Model description and simulation of present-day climate, *Max-Planck-Institut für Meteorologie, Tech. Rep.* 218, 1996.
- Roeckner, E., Bauml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M. A., Hagemann, S., Kirchner, I., Kornblüeh, L., Manzini, E., Rhodin, A., Schlese, U., Schulzweida, U., Tompkins, A. M.: The atmospheric general circulation model ECHAM5—Part I: Model description, *Max-Planck-Institut für Meteorologie, Tech. Rep.* 349, 2003.

- Roeckner, E., Brokopf, R., Esch, M., Giorgetta, M. A., Hagemann, S., Kornblueh, L., Manzini, E., Schlese, U., and Schulzweida, U.: Sensitivity of simulated climate to horizontal and vertical resolution in the ECHAM5 atmosphere model, *J. Climate*, 19, 3771–3791, doi:10.1175/JCLI3824.1, 2006.
- Roquet, F., Wunsch, C., and Madec, G.: On the patterns of wind-power input to the ocean circulation, *J. Phys. Oceanogr.*, 41, 2328–2342, doi:10.1175/JPO-D-11-024.1, 2011.
- Sandery, P.A., Brassington, G.B., Craig, A., and Pugh, T.: Impacts of Ocean–Atmosphere Coupling on Tropical Cyclone Intensity Change and Ocean Prediction in the Australian Region. *Mon. Weather Rev.*, 138, 2074–2091, doi:10.1175/2010MWR3101.1, 2010.
- Simmons A., Uppala S., Dee D., Kobayashi S.: ERA-Interim: New ECMWF reanalysis products from 1989 onwards, *ECMWF Newsletter*, 110, 26–35, 2006.
- Song, Y., Qiao, F., and Song, Z.: Ocean wave induced mixing effects on SST by cloud-radiation feedback in tropical atmosphere, *Geophysical Research Abstracts*, 14, 3982, 2012.
- Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., Salzmann, M., Schmidt, H., Bader, J., Block, K., Brokopf, R., Fast, I., Kinne, S., Kornblueh, L., Lohmann, U., Pincus, R., Reichler, T., Roeckner, E.: The atmospheric component of the MPI-M Earth system model: ECHAM6, *J. Adv. Model. Earth Sy.*, doi:10.1002/jame.20015, 2013.
- Sun F., Yu J.: Impacts of Central America gap winds on the SST annual cycle in the eastern Pacific warm pool, *Geophys. Res. Lett.*, 33, 2006.
- Saji, N. H., Goswami, B. N., Vinayachandran, P. N., and Yamagata, T.: A dipole mode in the tropical Indian Ocean, *Nature*, 401, 360–363, 1999.
- Sheffield, J., Barrett, A.P., Colle, B., Fernando, D. N., Fu, R., Geil, K.L., Hu, Q., Kinter, J., Kumar, S., Langenbrunner, B., Lombardo, K., Long, L.N., Maloney, E., Mariotti, A., Meyerson, J. E., Mo, K. C., Neelin, J. D., Nigam, S., Pan, Z., Ren, T., Ruiz-Barradas, A., Serra, Y. L., Seth, A., Thibeault, J. M., Stroeve, J. C., Yang, Z., and Yin, L.: North American Climate in CMIP5 Experiments. Part I: Evaluation of Historical Simulations of Continental and Regional Climatology, *J. Climate*, 26, 9209–9245, doi:10.1175/JCLI-D-12-00592.1, 2013.
- Simmons, A. J., Burridge, D. M., Jarraud, M., Girard, C., and Wergen, W.: The ECMWF medium-range prediction models development of the numerical formulations and the impact of increased resolution, *Meteorol. Atmos. Phys.*, 40, 28–60, 1989.
- Stevens, B., Giorgetta, M. A., Esch, M., Mauritsen, T., Crueger, T., Rast, S., Salzmann, M., Schmidt, H., Bader, J., Block, K., Brokopf, R., Fast, I., Kinne, S., Kornblueh, L., Lohmann, U., Pincus, R., Reichler, T., and Roeckner, E.: Atmospheric component of the MPI-M Earth System Model: ECHAM6, *J. Adv. Model. Earth Sy.*, 5, 146–172, doi: 10.1002/jame.20015, 2013.
- Storkey, D., Blockley, E. W., Furner, R., Lea, D. J., Martin, M. J., Barciela, R. M., Hines, A., Hyder, P., and Siddorn, J. R.: Forecasting the ocean state using NEMO: The new FOAM system, *J. Oper. Oceanogr.*, 3, 3–15, doi:10.1080/1755876X.2010.11020109, 2014.

- Tang, Y., Li, L., Dong, W., and Wang, B.: Reducing the climate shift in a new coupled model, *Chinese Sci. Bull.*, 61, 488-494, doi:10.1007/s11434-016-1033-y, 2016.
- Tett, S. F., Sherwin, T. J., Shrivastava, A., and Browne, O.: How Much Has the North Atlantic Ocean Overturning Circulation Changed in the Last 50 Years?, *J. Climate*, 27, 6325–6342, doi:10.1175/JCLI-D-12-00095.1, 2014.
- Thompson, D. W. J., Solomon, S.: Interpretation of recent Southern Hemisphere climate change, *Science*, 296, 895-899, 2002.
- Toyoda, T., Fujii, Y., Kuragano, T., Kamachi, M., Ishikawa, Y., Masuda, S., Sato, K., Awaji, T., Hernandez, F., Ferry, N., Guinehut, S., Martin, M. J., Peterson, K. A., Good, S. A., Valdivieso, M., Haines, K., Storto, A., Masina, S., Köhl, A., Zuo, H., Balmaseda, M., Yin, Y., Shi, L., Alves, O., Smith, G., Chang, Y.-S., Vernieres, G., Wang, X., Forget, G., Heimbach, P., Wang, O., Fukumori, I., and Lee, T., Intercomparison and validation of the mixed layer depth fields of global ocean syntheses, *Climate Dynamics*, 49: 753-773, doi:10.1007/s00382-015-2637-7, 2017.
- Valcke, S.: The OASIS3 coupler: a European climate modelling community software, *Geosci. Model Dev.*, 6, 373-388, doi:10.5194/gmd-6-373-2013, 2013.
- Vargashernandez, J. M., Wijffels, S., Meyers, G., Couto, A. B., and Holbrook, N. J.: Decadal characterization of Indo-Pacific Ocean subsurface temperature modes in SODA reanalysis, *J. Climate*, 28, 6113-6132, doi:10.1175/JCLI-D-14-00700.1, 2015.
- Wallace, J. M., and Thompson D. W. J.: The Pacific center of action of the Northern Hemisphere annular mode: Real or artefact?, *J. Climate*, 15, 1987-1991, 2002.
- Wang, C., Zhang, L., Lee, S., Wu, L., and Mechoso, C. R.: A global perspective on CMIP5 climate model biases, *Nat., Clim. Change*, 4, 201-205, doi:10.1038/nclimate2118, 2014.
- Wild, M., Folini, D., Schär, C., Loeb, N., Dutton, E. G., König-Langlo, G.: The global energy balance from a surface perspective, *Clim. Dynam.*, 40, 3107, doi:10.1007/s00382-012-1569-8, 2013.
- Wild, M., Folini, D., Hakuba, M. Z., Schar, C., Seneviratne, S. I., Kato, S., Rutan, D., Ammann, C., Wood, E. F., König-Langlo, G.: The energy balance over land and oceans: an assessment based on direct observations and CMIP5 climate models, *Clim. Dynam.*, 44, 3393-3429, doi:10.1007/s00382-014-2430-z, 2015.
- Wolff, J. O., Maier-Reimer, E., and Legutke, S.: The Hamburg Ocean Primitive Equation Model HOPE, German Climate Computer Center (DKRZ), Tech. Rep. 13, 98, 1997.
- Wu, R.G. and Kinter III, J. L.: Atmosphere-ocean relationship in the mid-latitude North Pacific: Seasonal dependence and east-west contrast, *J. Geophys. Res.*, 115, D06101, doi:10.1029/2009JD012579, 2010.
- Wu, L., and Liu, Z.: Decadal Variability in the North Pacific: The Eastern North Pacific Mode, *J. Climate*, 16, 3111-3131, doi:10.1175/1520-0442(2003)016<3111:DVITNP>2.0.CO;2, 2003.
- Xiao, C.: Application of a shape-persevering advection scheme in an OGCM and Coupled GCM (Master's thesis), Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China. (in Chinese), 2006.

- Xue, Y., Huang, B., Hu, Z., Kumar, A., Wen, C., Behringer, D. W., and Nadiga, S.: An assessment of oceanic variability in the NCEP climate forecast system reanalysis, *Clim. Dynam.*, 37, 2511-2539, doi:10.1007/s00382-010-0954-4, 2011.
- Ying, J., and Huang, P.: The Large-Scale Ocean Dynamical Effect on Uncertainty in the Tropical Pacific SST Warming Pattern in CMIP5 Models, *J. Climate*, 29, 8051–8065, doi:10.1175/JCLI-D-16-0318.1, 2016.
- Yu, J. Y., Kim, S. T.: Reversed spatial asymmetries between El Niño and La Niña and their linkage to decadal ENSO modulation in CMIP3 models, *J. Climate*, 24: 5423–5434, 2011.
- Zanchettin, D., Bothe, O., Graf, H. F., Lorenz, S. J., Luterbacher, J., Timmreck, C., and Jungclauss, J. H.: Background conditions influence the decadal climate response to strong volcanic eruptions, *J. Geophys. Res.-Atmos.*, 118, 4090–4106, doi:10.1002/jgrd.50229, 2013.
- Zhang, D., Blender, R., and Fraedrich, K.: Volcanoes and ENSO in millennium simulations: Global impacts and regional reconstructions in East Asia, *Theor. Appl. Climatol.*, 111, 437–454, doi:10.1007/s00704-012-0670-6, 2013.
- Zhang, L. and Zhao, C.: Processes and mechanisms for the model SST biases in the North Atlantic and North Pacific: A link with the Atlantic meridional overturning circulation, *J. Adv. Model. Earth Sy.*, 7, 739-758, doi:10.1002/2014MS000415, 2015.
- Zhang, X., Liang, S., Wild, M., and Jiang, B.: Analysis of surface incident shortwave radiation from four satellite products. *Remote Sens. Environ.*, 165, 186-202. doi:10.1016/j.rse.2015.05.015, 2015.
- Zhang, X., Liang, S., Wang, G., Yao, Y., Jiang, B., and Cheng, J.: Evaluation of the Reanalysis Surface Incident Shortwave Radiation Products from NCEP, ECMWF, GSFC, and JMA Using Satellite and Surface Observations, *Remote Sens.*, 8, 225, doi:10.3390/rs8030225, 2016.



**Table 1: Pattern Correlations of surface variables and SST deviation**

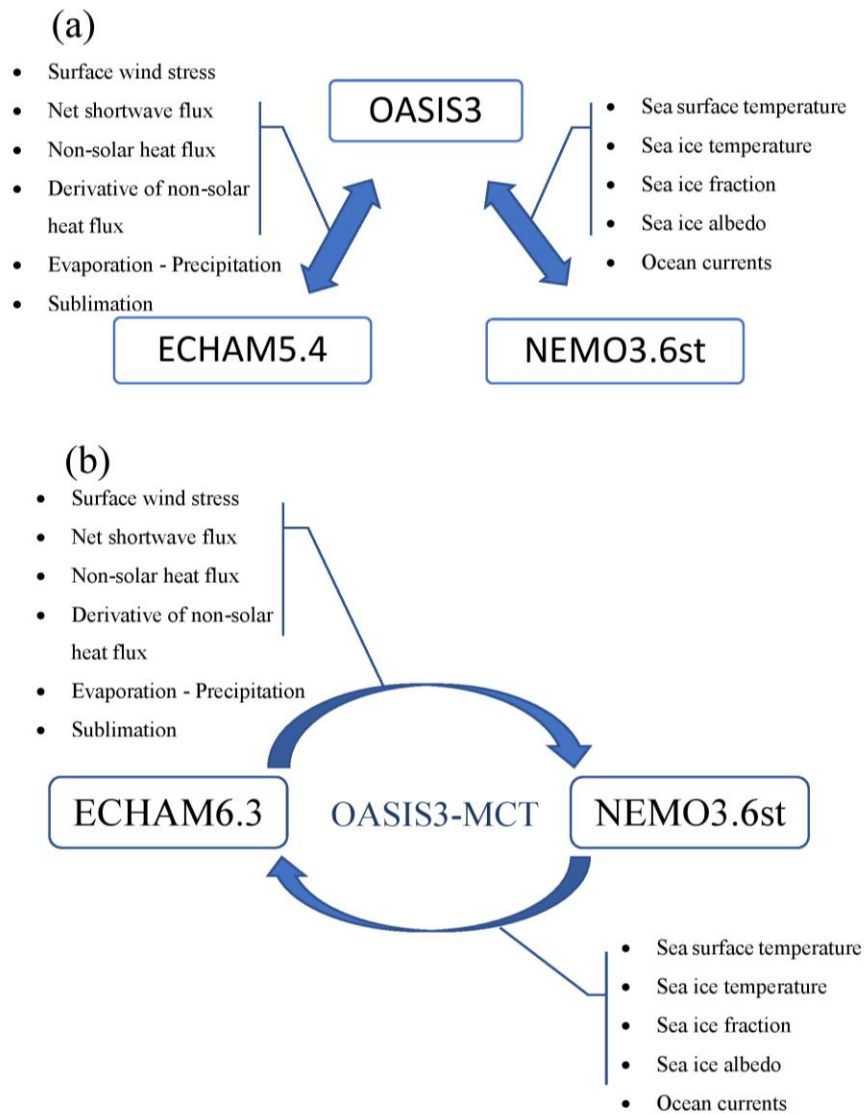
Model Replacement Variables	AGCM	OGCM
Surface zonal currents	0.006	-0.002
Surface meridional currents	0.0158	-0.018*
Sensible heat flux	-0.068*	-0.269*
Latent heat flux	-0.262*	<b>-0.393*</b>
Mean sea level pressure	-0.136*	-0.009
10m zonal wind	0.018*	-0.010
10m meridional wind	-0.025*	0.045*
Surface albedo	-0.101*	-0.022*
Net shortwave flux	0.192*	0.138*
Net longwave flux	<b>-0.269*</b>	-0.283*
Precipitation	0.100*	0.154*

Asterisk (\*) denotes the correlation coefficients passing the Student-t test above 99.9% confidence level. A larger number of grid points are involved that makes the threshold value relatively small. Numbers in boldface denotes the maximum absolute correlation value for the AGCM or OGCM replacement.

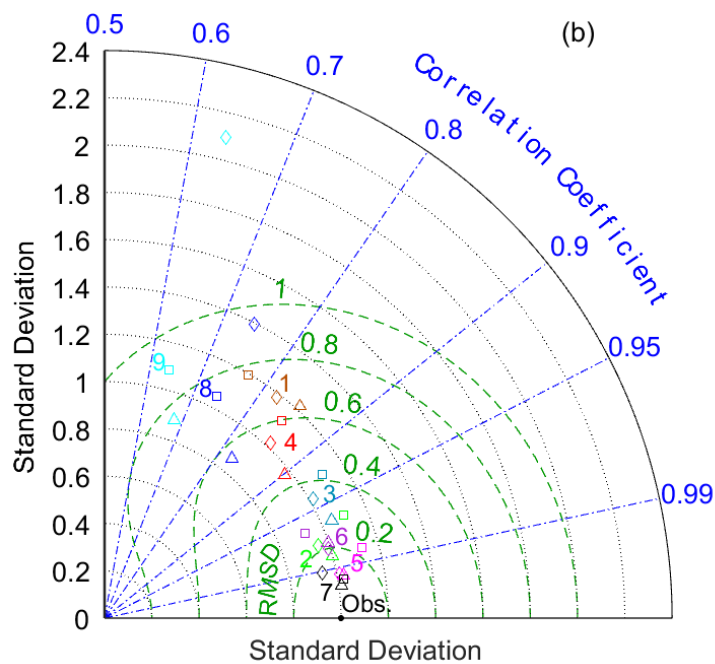
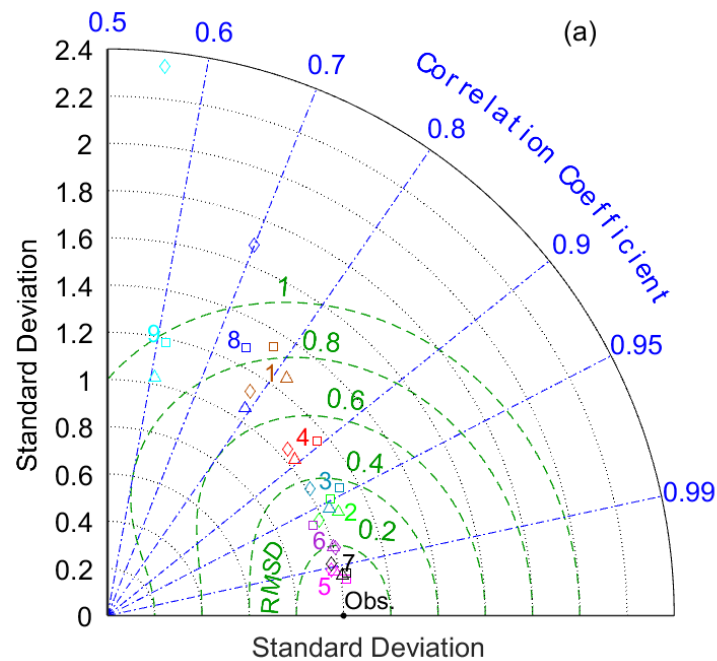
**Table 2: Pattern Correlations of surface variables and SST deviation over the Pacific**

Model Replacement Variables	AGCM	OGCM
Surface zonal currents	0.069*	-0.033*
Surface meridional currents	0.055*	-0.136*
Sensible heat flux	-0.286*	<b>-0.318*</b>
Latent heat flux	-0.302*	-0.315*
Mean sea level pressure	-0.246*	0.014
10m zonal wind	0.070*	0.039*
10m meridional wind	0.022	-0.005
Surface albedo	0.274*	0.034*
Net shortwave flux	<b>0.375*</b>	0.102*
Net longwave flux	-0.303*	-0.229*
Evaporation	-0.302*	-0.315*
Precipitation	0.103*	0.124*

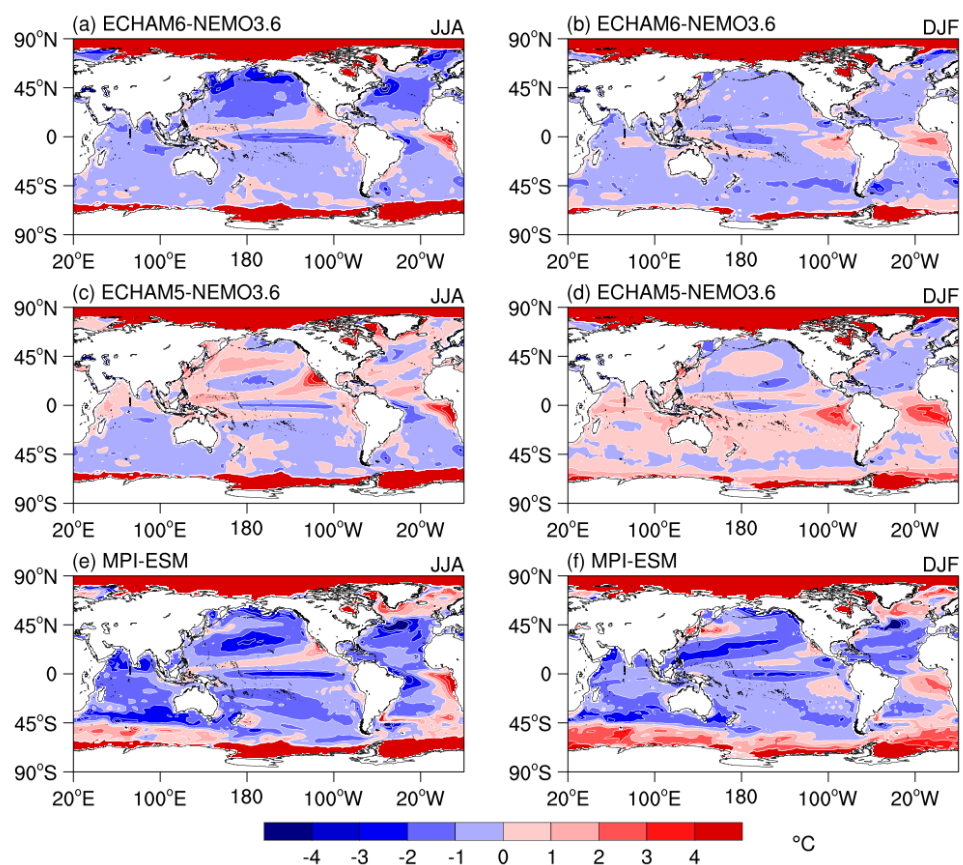
Asterisk (\*) denotes the correlation coefficients passing the Student-t test above 99.9% confidence level. A larger number of grid points are involved that makes the threshold value relatively small. Numbers in boldface denotes the maximum absolute correlation value for the AGCM or OGCM replacement.



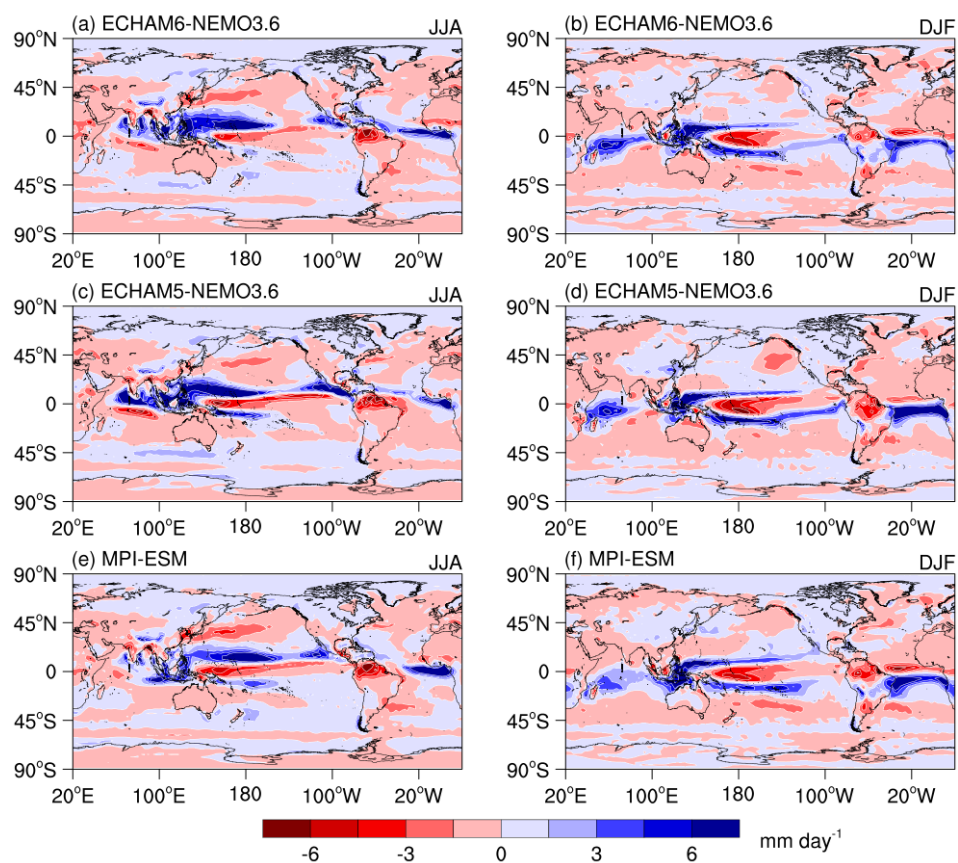
**Figure 1: Schematic structure of ECHAM5-NEMO3.6 (a), and ECHAM6-NEMO3.6 (b).**



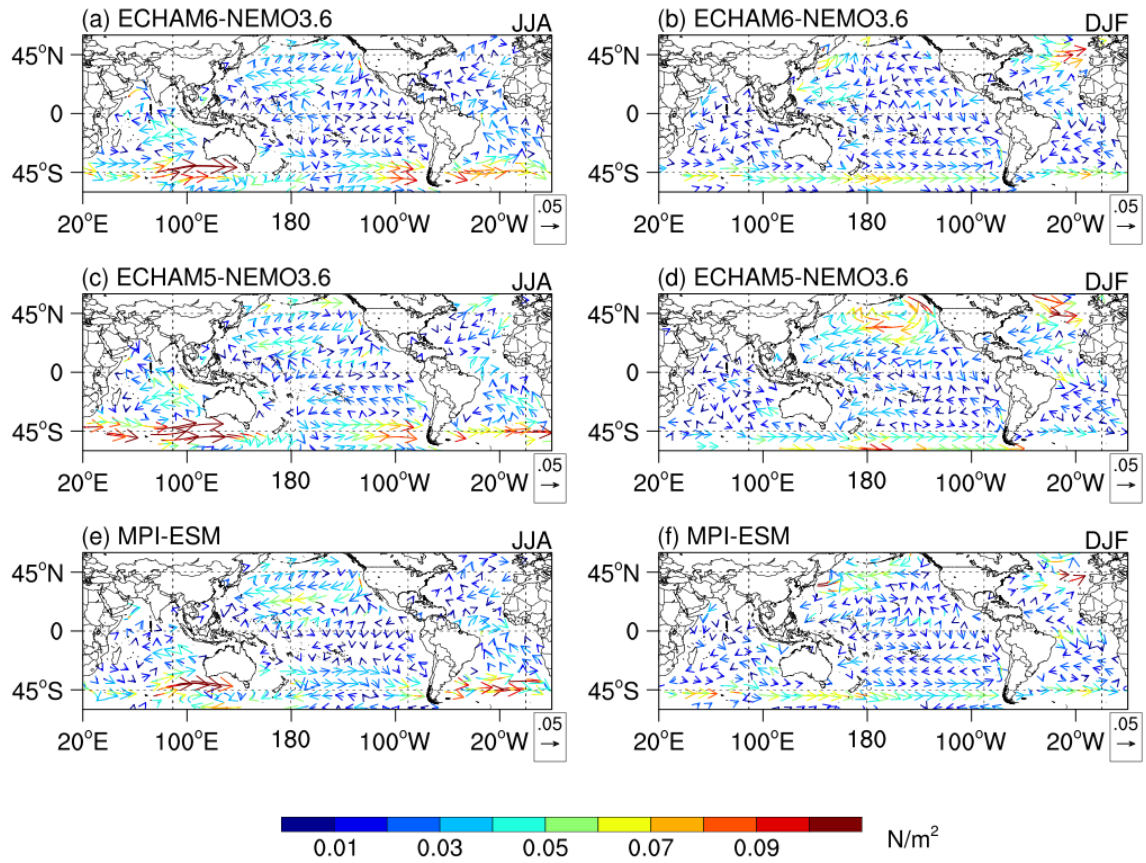
**Figure 2: Taylor diagram that exhibits a statistical comparison between the simulations and reanalysis data of nine selected variables in summer (a) and winter (b). Each number represents one variable: (1) precipitation, (2) mean sea level pressure, (3) zonal winds at 10m height, (4) meridional winds at 10m height, (5) 2m temperature, (6) total radiation flux (net shortwave plus net longwave), (7) SST, (8) sea surface zonal currents, (9) sea surface meridional currents. Upward-pointing triangles, squares and diamonds, respectively, represent the ECHAM6-NEMO3.6, ECHAM5-NEMO3.6, and the MPI-ESM results.**



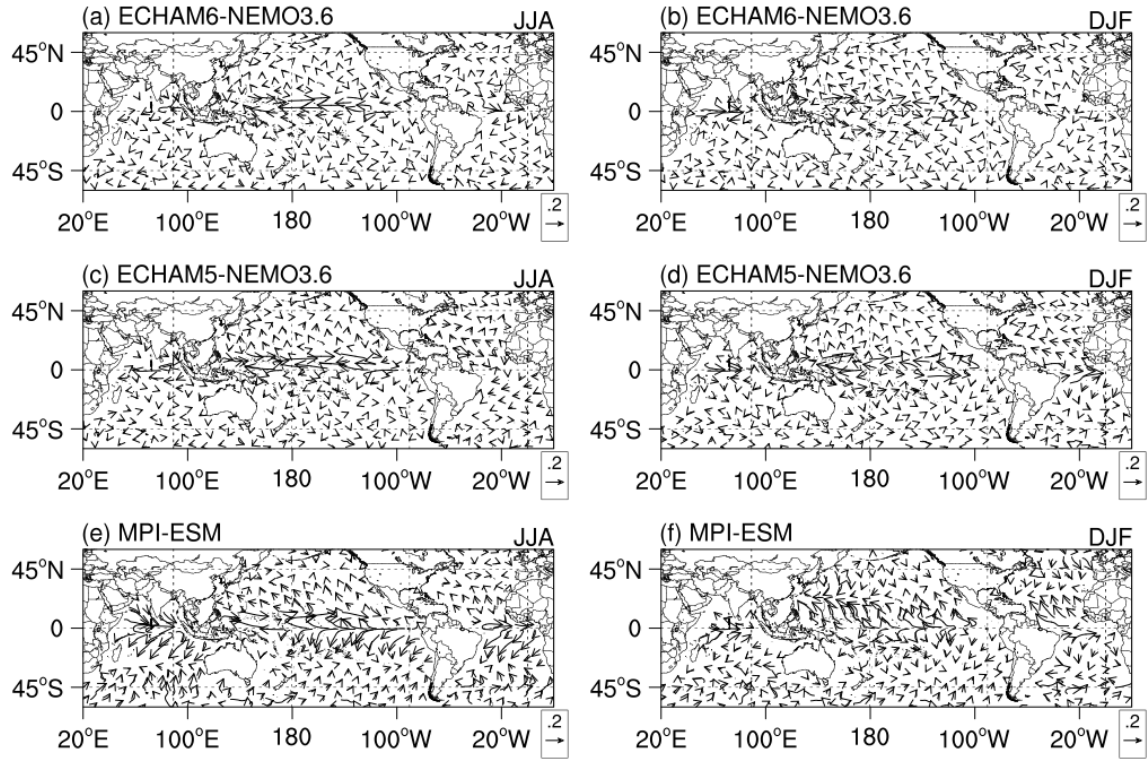
**Figure 3: Biases of the SST simulation in summer (left column) and winter (right column) corresponding to each CGCM: (a, b) ECHAM6-NEMO3.6, (c, d) ECHAM5-NEMO3.6, (e, f) MPI-ESM.**



**Figure 4: The same as Fig. 3 but for simulated precipitation climatology.**

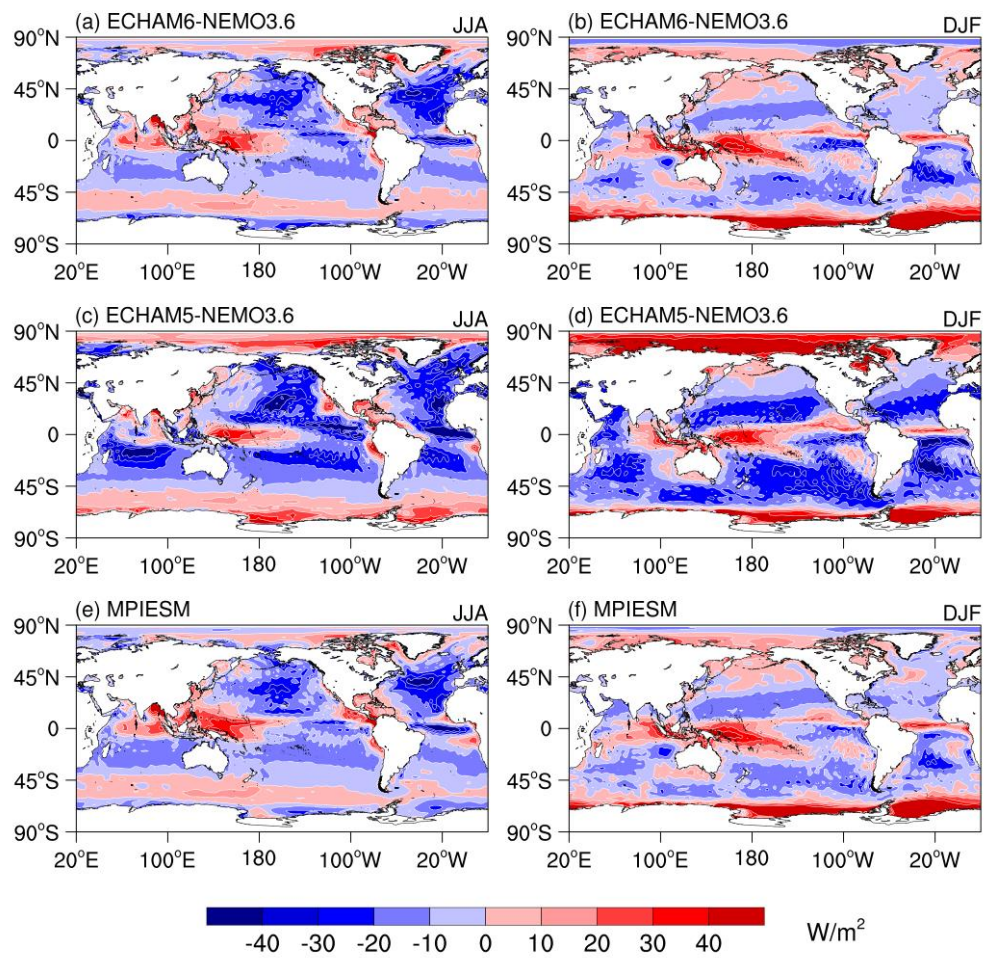


**Figure 5:** The same as Fig. 3 but for simulated surface wind stress (unit:  $\text{N/m}^2$ ).

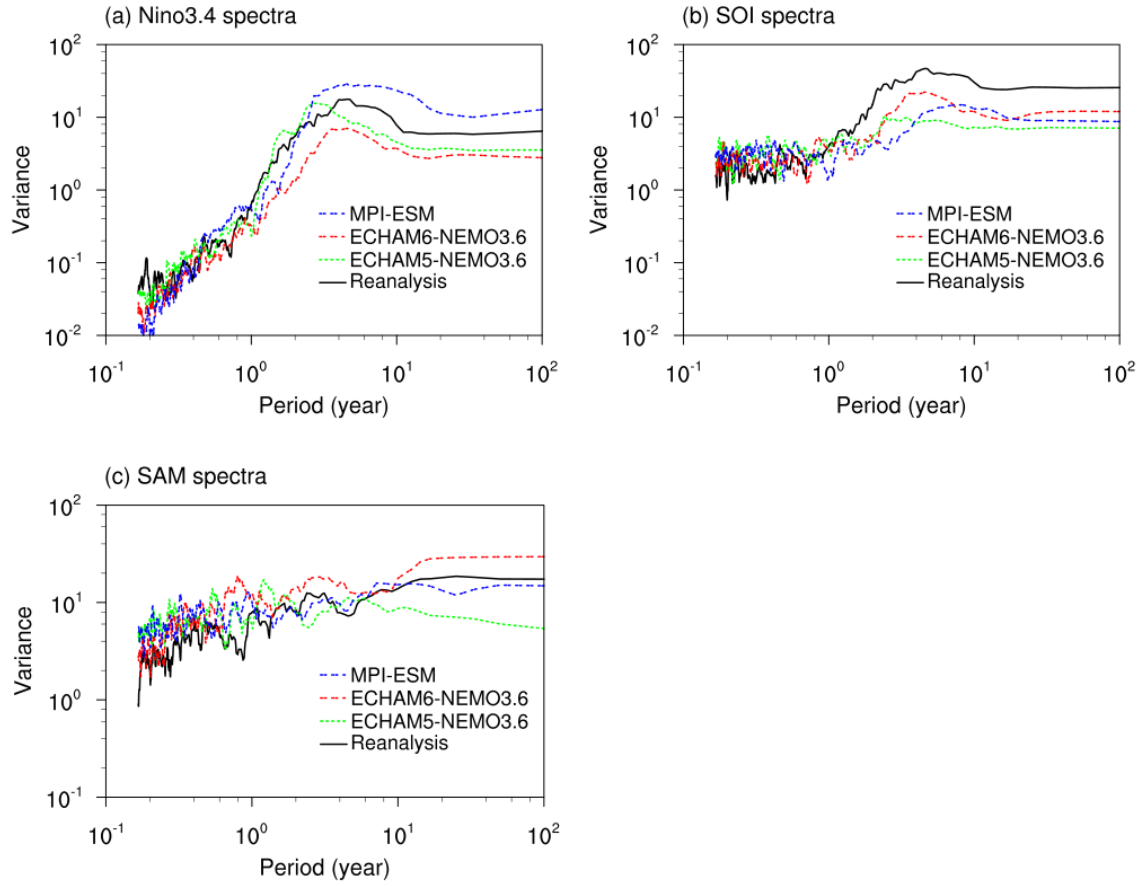


**Figure 6:** The same as Fig. 3 but for simulated surface currents (unit: m/s).

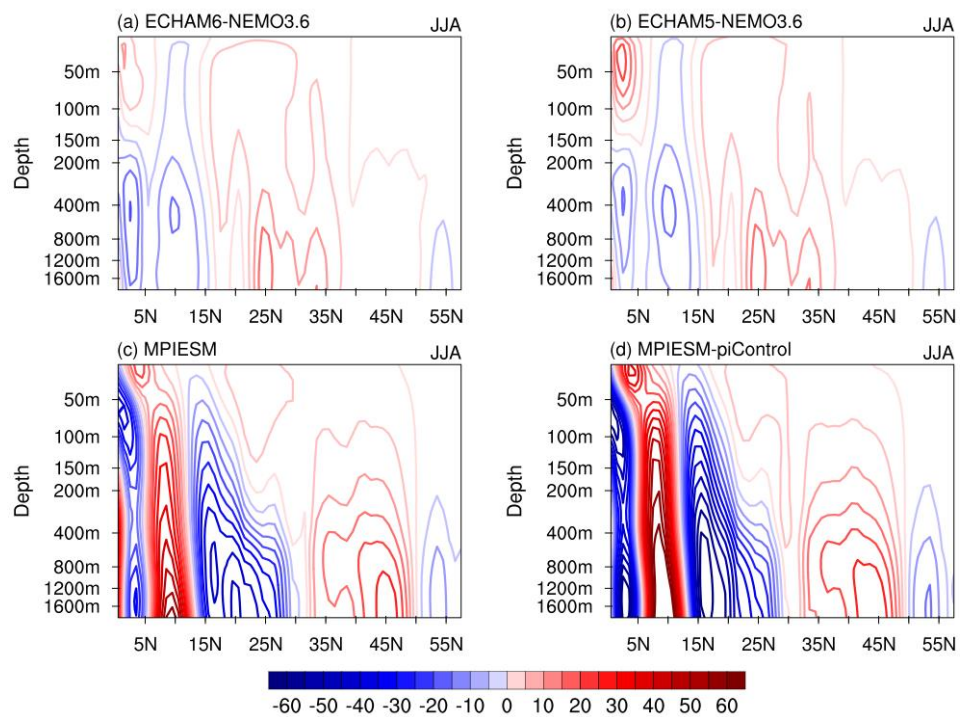




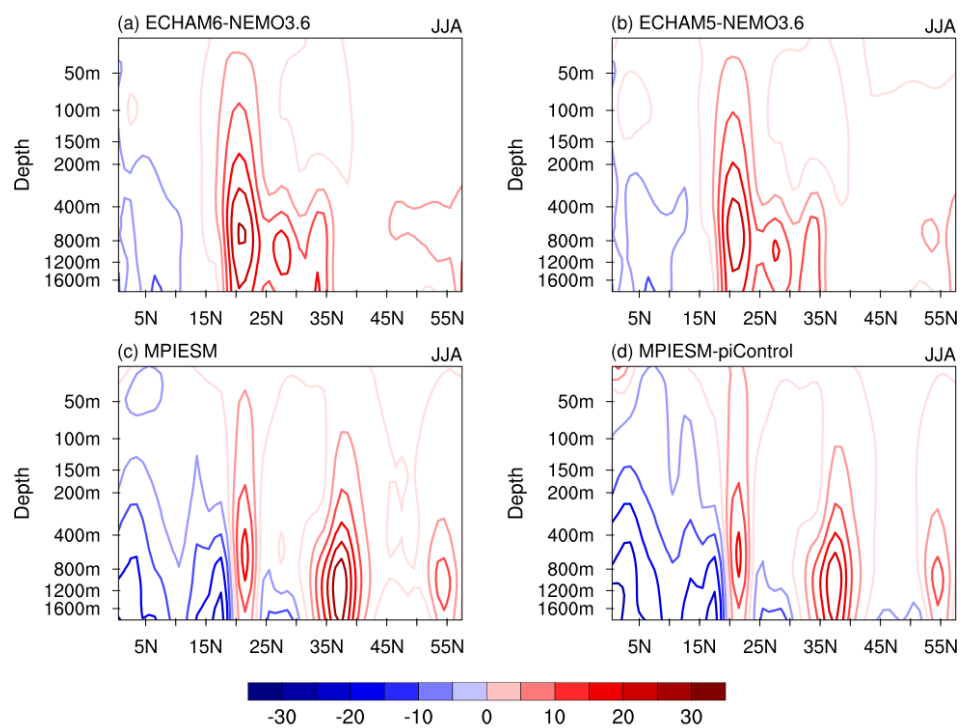
**Figure 7:** The same as Fig. 3 but for simulated total radiation.



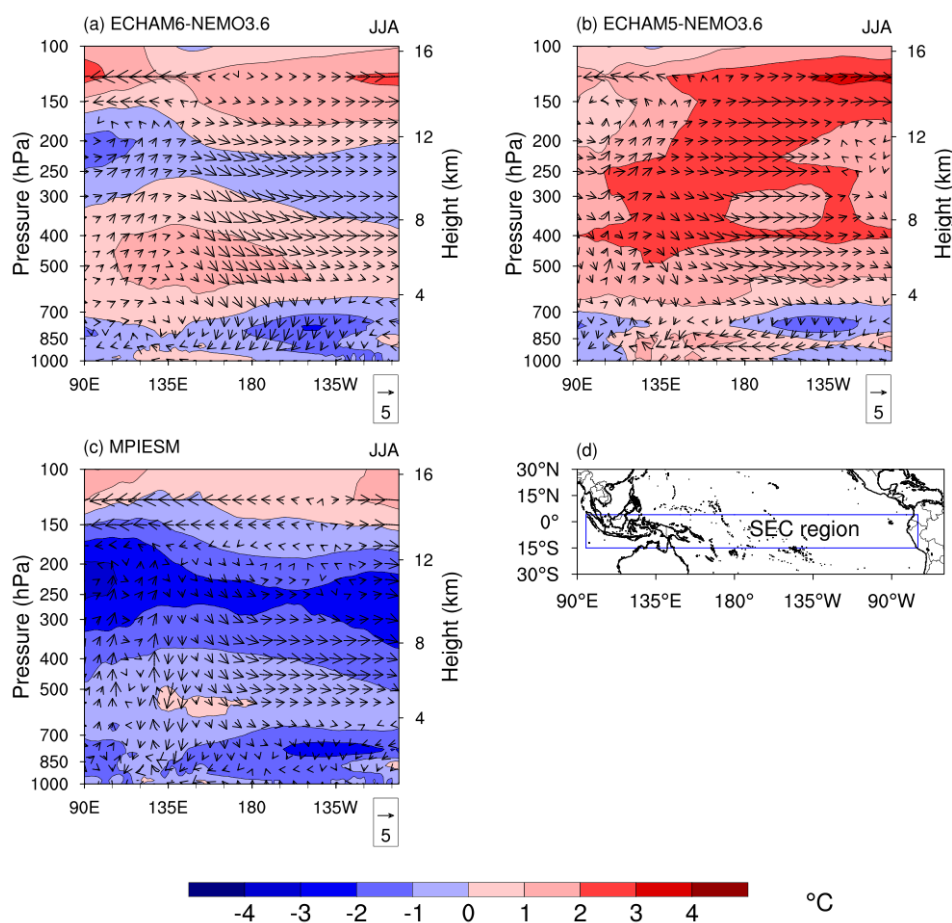
**Figure 8: Power spectra of (a) Niño3.4 index, (b) SOI, (c) SAM index. Solid line denotes the calculation results of reanalysis data, green dotted line denotes the ECHAM5-NEMO3.6 simulation, red dotted line denotes the ECHAM6-NEMO3.6 simulation, and blue dotted line denotes the MPI-ESM simulation.**



**Figure 9: Model biases of summer climatology of meridional overturning circulation simulation in North Pacific, (a) ECHAM6-NEMO3.6, (b) ECHAM5-NEMO3.6, (c) MPI-ESM, (d) MPI-ESM piControl, Unit: Sv.**



**Figure 10: Model biases of AMOC in summer, (a) ECHAM6-NEMO3.6, (b) ECHAM5-NEMO3.6, (c) MPI-ESM, (d) MPI-ESM piControl, Unit: Sv.**



**Figure 11: Model biases of the vertical structure of atmospheric circulation (vector) and the temperature (contour) over SEC area in summer for (a) ECHAM6-NEMO3.6, (b) ECHAM5-NEMO3.6, (c) MPI-ESM. The SEC region is marked as the blue box in (d). Vector units: m/s, contour unit: °C.**

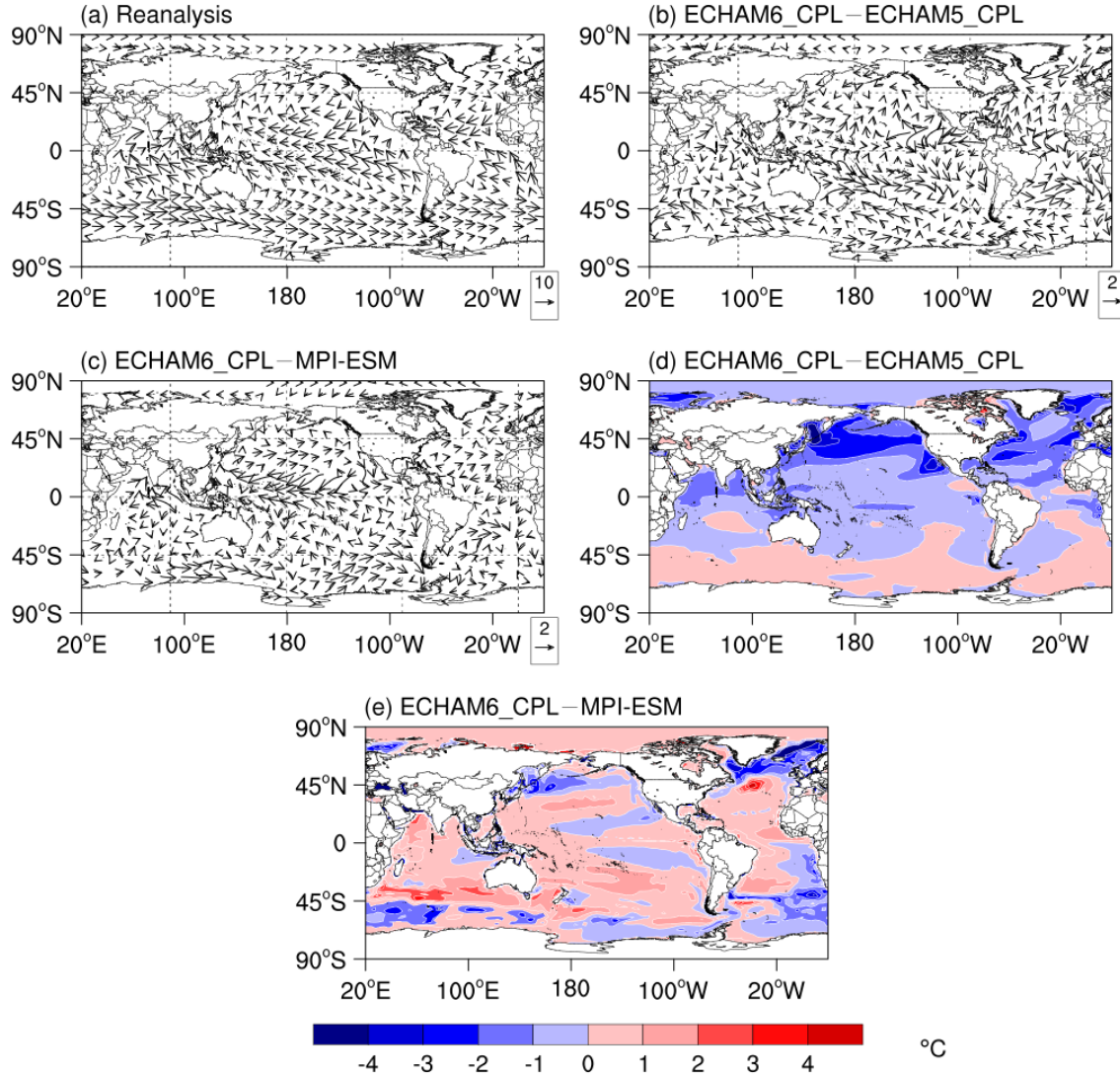
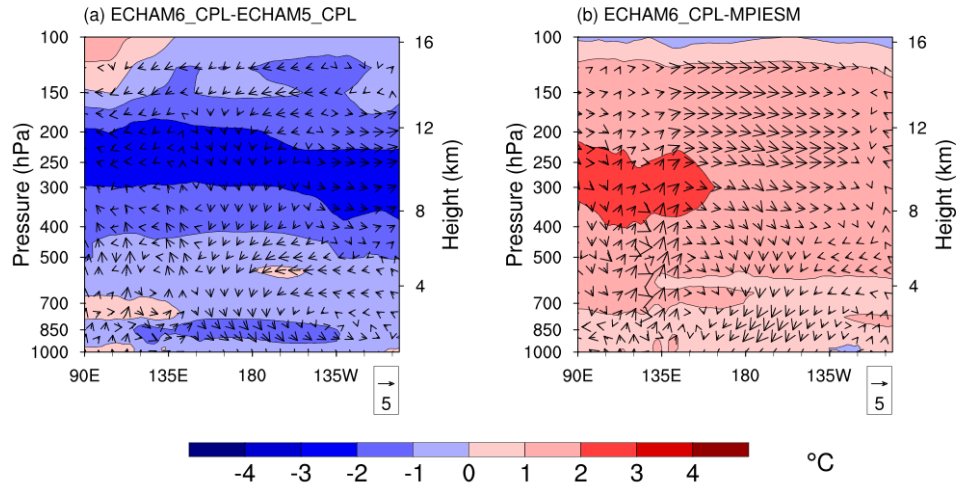
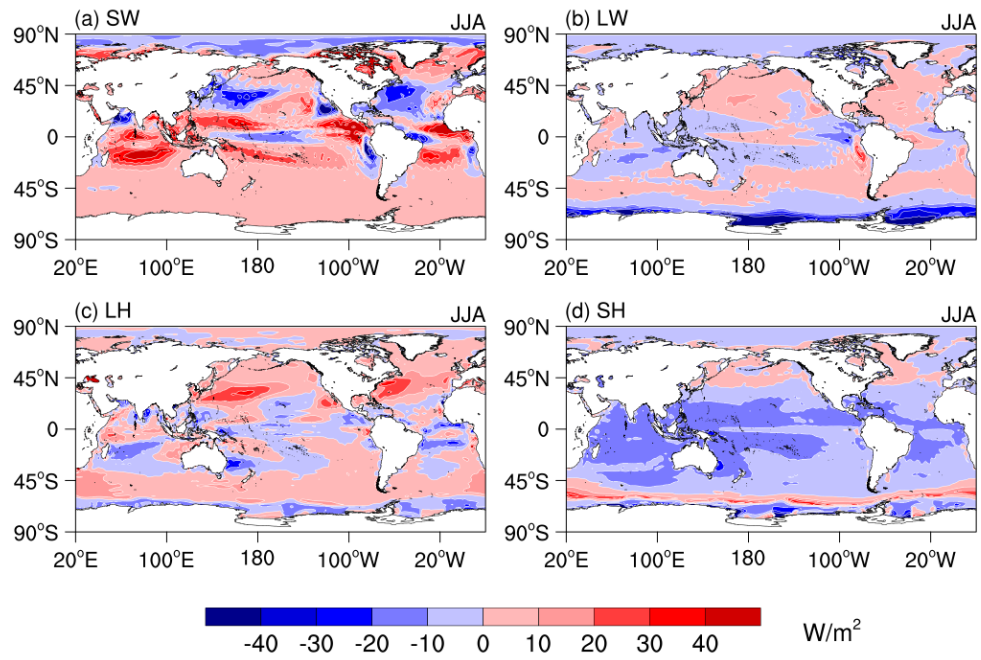


Figure 12: Summer climatology of 10m wind for (a) reanalysis data, (b) model differences between ECHAM6-NEMO3.6 and ECHAM5-NEMO3.6, (c) model differences between ECHAM6-NEMO3.6 and MPI-ESM. SST simulation differences for (d) between ECHAM6-NEMO3.6 and ECHAM5-NEMO3.6, and (e) between ECHAM6-NEMO3.6 and MPI-ESM. **Vector units: m/s, contour unit: °C.**



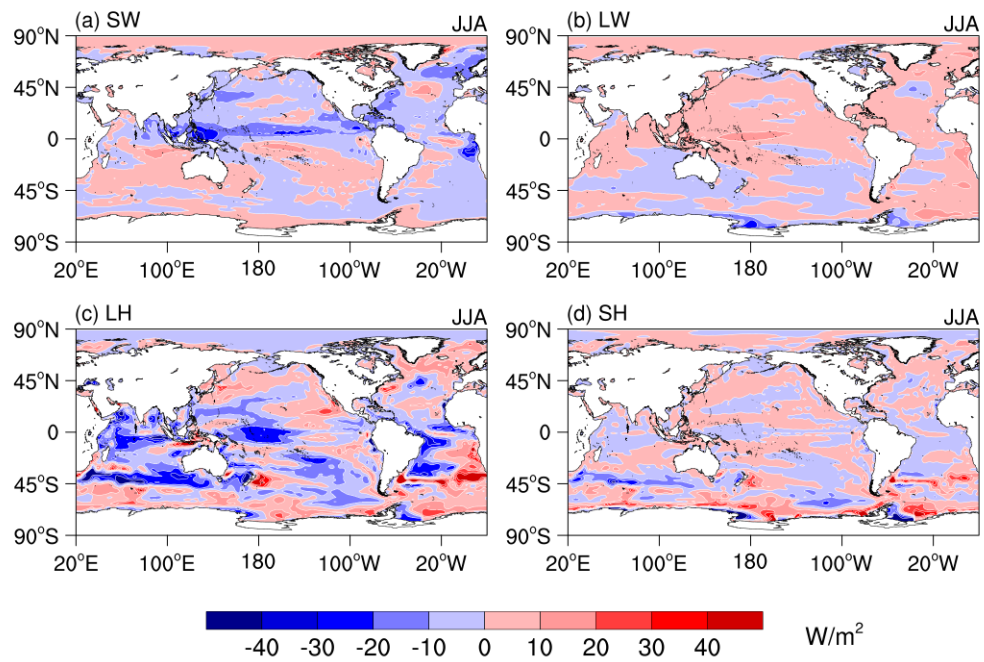


**Figure 13: Simulation differences in the vertical structure of atmospheric circulation and the temperature over the SEC area: (a) between ECHAM6-NEMO3.6 and ECHAM5-NEMO3.6, (b) between ECHAM6-NEMO3.6 and MPI-ESM. Vector units: m/s, contour unit: °C.**



**Figure 14: Simulation differences in radiation budget between ECHAM6-NEMO3.6 and ECHAM5-NEMO3.6.**





**Figure 15: Simulation differences in radiation budget between ECHAM6-NEMO3.6 and MPI-ESM.**