



Systematic improvement in simulated latent and sensible heat fluxes over tropical oceans in AMIP6 models compared to AMIP5 models with the same resolutions

Xin Zhou^{a,*}, Pallav Ray^a, Bradford S. Barrett^{b,c}, Pang-Chi Hsu^d

^a Florida Institute of Technology, Melbourne, FL, United States of America

^b United States Naval Academy, Annapolis, MD, United States of America

^c Air Force Office of Scientific Research, Santiago, Chile

^d Nanjing University of Information Science & Technology, Nanjing, China



ARTICLE INFO

Keywords:

Latent heat flux
Sensible heat flux
AMIP5
AMIP6
OAFlux
Tropical Oceans

ABSTRACT

Recent studies have shown that most Atmospheric Model Intercomparison Project (AMIP5) models that participated in the fifth phase of the coupled model intercomparison project (CMIP5) overestimate surface latent (Q_{LH}) and sensible (Q_{SH}) heat fluxes over tropical oceans. This study aims to quantify and understand any improvement in simulated Q_{LH} and Q_{SH} in AMIP6 models compared to AMIP5 models with the same horizontal resolutions. To accomplish this, we compare the Ocean Atmosphere air-sea Flux (OAFlux) dataset to nine AMIP5 and nine AMIP6 models integrated over 30 years (1979–2008). We found that all the AMIP5 and AMIP6 models overestimate both components of fluxes, but all of the AMIP6 models performed better than their AMIP5 counterparts over the tropical oceans (30°S–30°N). Yet, systematic spatial biases remain, which leads to only a small improvement in the simulated fluxes. For example, the root mean squared error (RMSE) and mean bias in Q_{LH} in AMIP5 are 30 and 21 W m^{-2} compared to 28 and 19 W m^{-2} in AMIP6. This is an improvement of 2 W m^{-2} for both RMSE and bias in Q_{LH} , yielding a reduction in Q_{LH} bias by 10% and RMSE by 7%. For Q_{SH} , an improvement of 1 W m^{-2} is seen in both bias and RMSE, yielding a reduction in Q_{SH} bias by 25% and RMSE by 13%. The primary reason for the improvement in Q_{LH} in AMIP6 is the better representation of the 10m wind speed and air-sea humidity difference than those in AMIP5. The improvement in Q_{SH} is due to an improvement in air-sea temperature difference. These results offer guidance to the modeling community to improve model simulation of near-surface meteorological variables in the tropics.

1. Introduction

It is well known that the ocean and atmosphere interact through air-sea fluxes of mass, energy, and momentum. These fluxes play important roles in the ocean's influence on weather and climate and the atmosphere's influence on ocean variability (e.g., [Jiang et al., 2005](#); [DeMott et al., 2015](#); [Hagos et al., 2021](#)). The net surface heat flux includes two turbulent terms (latent heat flux Q_{LH} , and sensible heat flux Q_{SH}) and two radiation terms (shortwave and longwave). The bulk formulae to calculate latent and sensible heat fluxes are given by:

$$Q_{LH} = \rho L_E V_a C_E (q_s - q_a) \quad (1)$$

$$Q_{SH} = \rho C_p V_a C_H (T_s - T_a) \quad (2)$$

where ρ is the density of air at the air-sea surface determined using the ideal gas law, L_E is the latent heat of evaporation ($\sim 2.5 \times 10^6 \text{ J Kg}^{-1}$), V_a is the wind speed at 10 m, C_E and C_H are the bulk transfer coefficients of latent heat and sensible heat, respectively; q_s is the saturation specific humidity at the sea surface temperature (SST), q_a is the specific humidity of near-surface air, C_p is the isobaric specific heat of air ($1008 \text{ J Kg}^{-1} \text{ K}^{-1}$), T_s is the SST, and T_a is the near-surface air temperature.

The Q_{LH} and Q_{SH} can significantly influence the oceanic mixed layer and the atmospheric planetary boundary layer (PBL) (e.g., [Sengupta et al., 2002](#); [DeMott et al., 2014](#)). As a result, evaluation of these flux components in contemporary models is important. For example, [Zhang et al. \(2018a, 2018b\)](#) assessed the performance of the multi-model ensemble (MME) based on models included in the fifth phase of the

* Corresponding author at: Florida Institute of Technology, Melbourne, FL 32901, USA.

E-mail address: xzhou2013@my.fit.edu (X. Zhou).

Table 1

Description of the nine AMIP5 and AMIP6 models used in this study.

AMIP5 Model (resolution, vertical levels)	AMIP6 Model (resolution, vertical levels)	Institution	Country	Reference
1 ACCESS1-0 ($1.25^\circ \times 1.9^\circ$, 38)	ACCESS-ESM1-5 ($1.25^\circ \times 1.9^\circ$, 38)	Collaboration for Australian Weather and Climate Research	Australia	Ziehn et al. (2020)
2 BNU-ESM ($2.8^\circ \times 2.8^\circ$, 26)	BCC-ESM1 ($2.8^\circ \times 2.8^\circ$, 26)	Beijing Climate Center	China	Wu et al. (2020)
3 CanAM4 ($2.8^\circ \times 2.8^\circ$, 26)	CanESM5 ($2.8^\circ \times 2.8^\circ$, 49)	Canadian Centre for Climate Modeling and Analysis	Canada	Swart et al. (2019)
4 CESM1-CAM5 ($0.94^\circ \times 1.25^\circ$, 27)	CESM2 ($0.94^\circ \times 1.25^\circ$, 32)	National Science Foundation Department of Energy, National Center for Atmospheric Research	USA	Danabasoglu et al. (2020)
5 GISS-E2-R ($2^\circ \times 2.5^\circ$, 40)	GISS-E2-1-G ($2^\circ \times 2.5^\circ$, 40)	NASA Goddard Institute for Space Studies	USA	Bauer et al. (2020)
6 INM-CM4 ($1.5^\circ \times 2^\circ$, 21)	INM-CM4-8 ($1.5^\circ \times 2^\circ$, 21)	Institute for Numerical Mathematics	Russia	Volodin et al. (2019)
7 MIROC5 ($1.4^\circ \times 1.4^\circ$, 40)	MIROC6 ($1.4^\circ \times 1.4^\circ$, 81)	Atmosphere and Ocean Research Institute, National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	Japan	Shiogama et al. (2019)
8 MPI-ESM-LR ($1.9^\circ \times 1.9^\circ$, 47)	MPI-ESM1-2-LR ($1.9^\circ \times 1.9^\circ$, 47)	Max Planck Institute for Meteorology	Germany	Mauritsen et al. (2019)
9 MRI-CGCM3 ($1.1^\circ \times 1.1^\circ$, 48)	MRI-ESM2-0 ($1.1^\circ \times 1.1^\circ$, 80)	Meteorological Research Institute	Japan	Yukimoto et al. (2019)

Coupled Model Intercomparison Project (CMIP5). They found that the mean Q_{LH} in the MME compared well with observations except near coastal regions and over the tropical oceans; in those areas, the MME overestimated the Q_{LH} . Cao et al. (2015) evaluated 14 CMIP5 models for their accuracy in capturing Q_{LH} in the Pacific. They found that models are able to capture the climatological distribution of latent heat flux reasonably well, but the amplitudes are generally overestimated. Similar model evaluation studies have also been conducted using coupled models for land surface energy and water budgets (e.g., Schwingsackl et al., 2018; Yang et al., 2019; Li et al., 2021).

Most model-data comparison studies concerning surface heat fluxes evaluate coupled ocean-atmosphere models (e.g., Zhang et al., 2018a, 2018b; Wild, 2020; Li et al., 2021). Therefore, we now have a better understanding of the performance of the coupled models in simulation of Q_{LH} and Q_{SH} . However, model evaluation studies for surface heat fluxes in atmospheric-only models are rare (Zhou et al., 2019, 2020). As a result, we have limited understanding of the accuracy of Q_{LH} and Q_{SH}

in atmosphere-only models. Moreover, given that the atmospheric model intercomparison project (AMIP) models from the 5th (AMIP5) and 6th (AMIP6) phases of CMIP are forced with prescribed SST, the intermodel differences coming from the differences in SST are negligible. Therefore, any differences in Q_{LH} and Q_{SH} in AMIP models arguably come from other parameters used to calculate these fluxes. On the other hand, the SST difference between CMIP models can be significant (Li and Xie, 2012), leading to intermodel differences in simulated fluxes (e.g., Reichler and Kim, 2008; Li and Xie, 2012).

Moreover, understanding the role of different variables in the AMIP simulations may help in understanding the behavior of their coupled counterparts. For example, Hourdin et al. (2015) compared 20 CMIP5 models and stand-alone AMIP atmospheric simulations with prescribed SST. They found that surface evaporation controls the amplitude of the surface temperature response to the overestimation of surface heat flux. DeMott et al. (2014) analyzed three general circulation models and compared them to their atmosphere-only counterparts. They showed that coupled models exhibit enhanced surface fluxes and intraseasonal oscillation than atmosphere-only models. They also found that specific humidity can explain ~20% of the variance in tropical Indian Ocean latent heat flux variance, and temperature can explain 50% of tropical Indian Ocean sensible heat flux. Xiang et al. (2017) found that the spurious double intertropical convergence zone (double ITCZ) in CMIP5 coupled models can be related to their atmospheric components.

Among the few studies that explored the biases concerning Q_{LH} in coupled and uncoupled models, Găinușă-Bogdan et al. (2018) found a systematic spatial bias pattern in Q_{LH} that was due to bias in near-surface winds and relative humidity in CMIP5 models. Zhou et al. (2020) found that most AMIP5 models overestimate Q_{LH} and Q_{SH} over oceans away from the equator, especially in the northern and southern Pacific Ocean and the southern Indian Ocean. More precisely, Zhou et al. (2020) found that all 20 AMIP5 models overestimated Q_{LH} , and 18 out of 20 AMIP5 models overestimated Q_{SH} . We further extend Zhou et al. (2020) with the following goals in mind: (i) to quantify and understand any improvement in the simulated Q_{LH} and Q_{SH} in AMIP6 compared to AMIP5 that is independent of model resolution (we use nine different AMIP models that have the same resolution in versions 5 and 6); (ii) to find the parameters that are responsible for biases in these models; (iii) to find out which parameters led to improvement, if any, in simulated fluxes in AMIP6 compared with AMIP5; and (iv) to explore the effect of intermodel resolution differences on simulated fluxes.

The rest of the paper is organized as follows: A description of the AMIP5 and AMIP6 models, observational data, and methods are given in section 2, followed by detailed model-data comparisons and the possible causes behind model bias in section 3. Conclusions are given in section 4.

2. Model, data, and method

2.1. AMIP models

The AMIP project was initiated in 1989 to generate and provide systematic evaluations of comparable integrations from the atmospheric GCMs (AGCMs) (Gates, 1992, 1999). AMIP5 (Taylor, 2001) and AMIP6 (Eyring et al., 2016; Feng et al., 2020) are included as an integral part of CMIP5 and CMIP6 to evaluate the performance of AGCMs. As an extension of AMIP5, several improvements were made in AMIP6. For example, AMIP6 models used the latest dataset that is consistent with the reanalysis to overcome the large and unrealistic changes in the sea ice concentration (Hurrell et al., 2008; Eyring et al., 2016). For the SST, both AMIP5 and AMIP6 models have valid data defined at all points to remove the SST bias caused by the undefined points (Hyder et al., 2018). In addition, identical observed SST and sea ice concentrations are used in all simulations in the AMIP6 experimental protocol (Eyring et al., 2016).

Out of the 20 AMIP5 models studied in Zhou et al. (2020), only nine

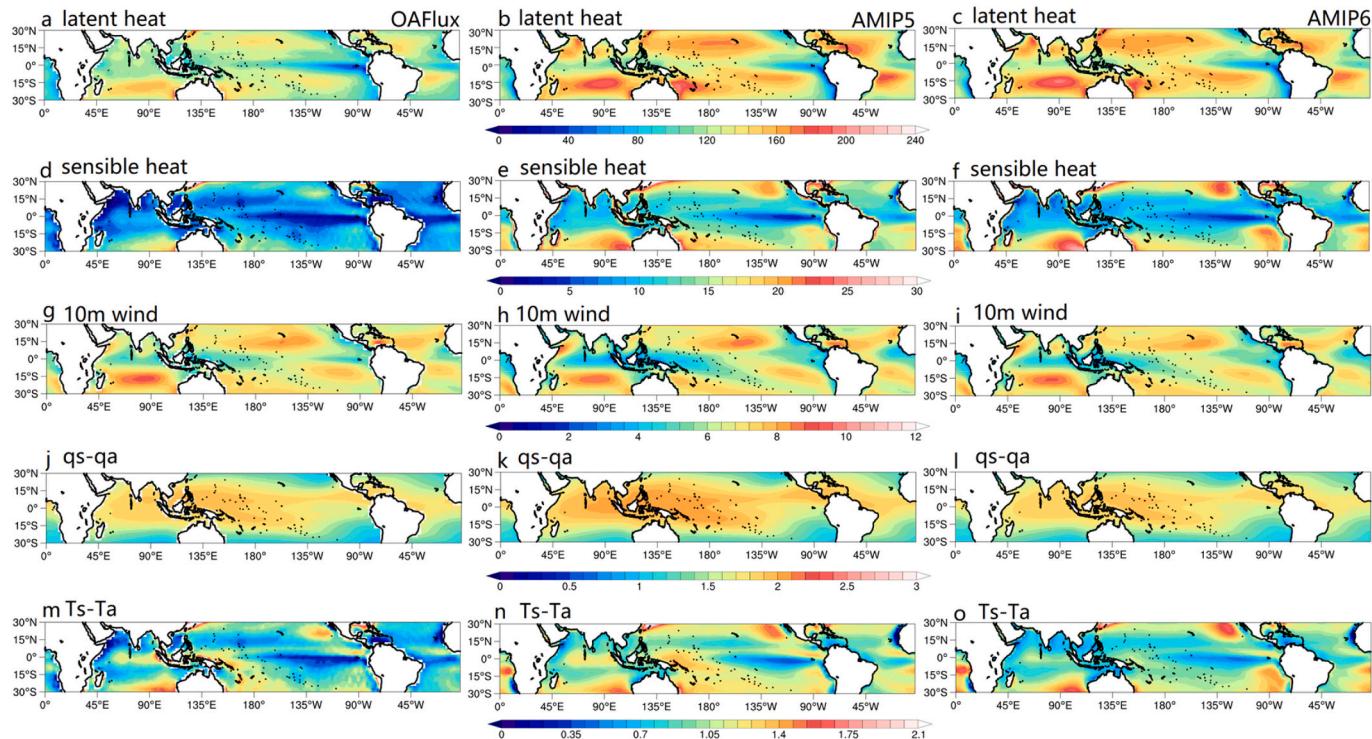


Fig. 1. Annual average (1979–2008) of the turbulent heat fluxes (W m^{-2}) and bulk variables: 10 m wind (m s^{-1}), $q_s - q_a$ (g kg^{-1}) and $T_s - T_a$ (K) from (left panel) OAFlux, (middle panel) AMIP5, and (right panel) AMIP6 over the tropical oceans.

were used here because (i) these nine models with varying resolutions participated in AMIP6 without change in horizontal resolutions (the horizontal resolution is maintained from AMIP5; Table 1), and (ii) all parameters needed for this study were available. The horizontal resolutions of these models range from $0.94^\circ \times 1.25^\circ$ (CESM1-CAM5 and CESM2) to $2.8^\circ \times 2.8^\circ$ (BNU-ESM, CanAM4, BCC-ESM1, and CanESM5). The number of vertical layers varies from 21 (INM-CM4) to 48 (MRI-CGCM3) in AMIP5 models, and from 21 (INM-CM4-8) to 81 (MIROC6) in AMIP6 models. Out of these 9 models, 4 AMIP6 models (CanESM5, CESM2, MIROC6, MRI-ESM2-0) have more vertical levels than their AMIP5 version. All calculations were conducted for simulations spanning 30 years (1979–2008). All the models used observed monthly SSTs (Fiorino, 1996; Hurrell et al., 2008; Li et al., 2015; Eyring et al., 2016) as the lower boundary conditions.

2.2. OAFlux data

Monthly latent and sensible heat fluxes, and relevant bulk variables (V_a, T_s, T_a, q_a , see equations 1 and 2) were taken from the Objectively Analyzed air-sea heat Fluxes (OAFlux; $1^\circ \times 1^\circ$; Yu and Weller, 2007; Yu et al., 2008) over the same period (1979–2008). The other popular blended flux product, TROPFlux (Kumar et al., 2012), has also been found to perform on par with the OAFlux when compared with observations (de Szeke et al., 2015). However, a few other studies (e.g., Rahaman and Ravichandran, 2013) found OAFlux to perform better than TROPFlux for Q_{LH} and specific humidity over the North Indian Ocean when compared to in situ observations. Kumar et al. (2012) found that OAFlux is better than TROPFlux, especially over regions with abundant low clouds, such as the eastern equatorial Pacific and Atlantic oceans. Pinker et al. (2014) showed that OAFlux biases in the monthly latent heat and sensible heat fluxes at the surface are less than 5 W m^{-2} and 2 W m^{-2} , respectively, and are expected to be much smaller for seasonal and annual means and when averaged over a larger region. A detailed assessment of the accuracy latent and sensible heat fluxes in OAFlux over the global oceans is provided in Bentamy et al. (2017). For

consistent model-data comparisons, we interpolate all model output and observational data to a uniform resolution of $2.5^\circ \times 2.5^\circ$. Also, for simplicity, we refer to OAFlux as an observation.

3. Methods

To find the bulk factors that are responsible for the structure of spatial bias in latent and sensible heat fluxes, we apply the decomposition method as in Hourdin et al. (2015), which we describe here briefly.

The Q_{LH} can be expressed as:

$$Q_{LH} = \gamma |V_a| q_{sat}(T_a) [\alpha L / R_v T_a^2 \delta T + (\alpha - RH)] \quad (3)$$

where $\gamma = \rho L_E C_E$, R_v is the gas constant for water vapor, and RH is the relative humidity, $\delta T = T_s - T_a$, and $\alpha = 0.98$, which accounts for the smaller evaporation of salty water than fresh water.

The bias in Q_{LH} is computed as:

$$\Delta Q_{LH} = \Delta L H_{dyn} + \Delta L H_{Qsat} + \Delta L H_{RH} + \Delta L H_{\delta T} \quad (4)$$

where $\Delta L H_{dyn}$ is contribution of bias from V_a , $\Delta L H_{Qsat}$ is contribution of bias from saturation humidity (Q_{sat}), $\Delta L H_{RH}$ is contribution of bias from relative humidity (RH), and $\Delta L H_{\delta T}$ is contribution of bias from δT , and they are computed as:

$$\Delta L H_{dyn} = \gamma \Delta |V_a| q_{sat}(T_a) \left[\frac{\alpha L}{R_v} T_a^2 \delta T + (\alpha - RH) \right] \quad (5)$$

$$\Delta L H_{Qsat} = \gamma |V_a| q_{sat}(T_a) L / R_v T_a^2 [\alpha \delta T (L / R_v T_a^2 - 2 / T_a) + (\alpha - RH)] \Delta T_a \quad (6)$$

$$\Delta L H_{RH} = -\gamma |V_a| q_{sat}(T_a) \Delta RH \quad (7)$$

$$\Delta L H_{\delta T} = \gamma |V_a| \alpha \beta q_{sat}(T_a) \Delta \delta T \quad (8)$$

Where $\beta = \frac{L}{R_v T_a^2} = 0.06$ is a constant.

q_{sat} was calculated as (Smith et al., 1999):

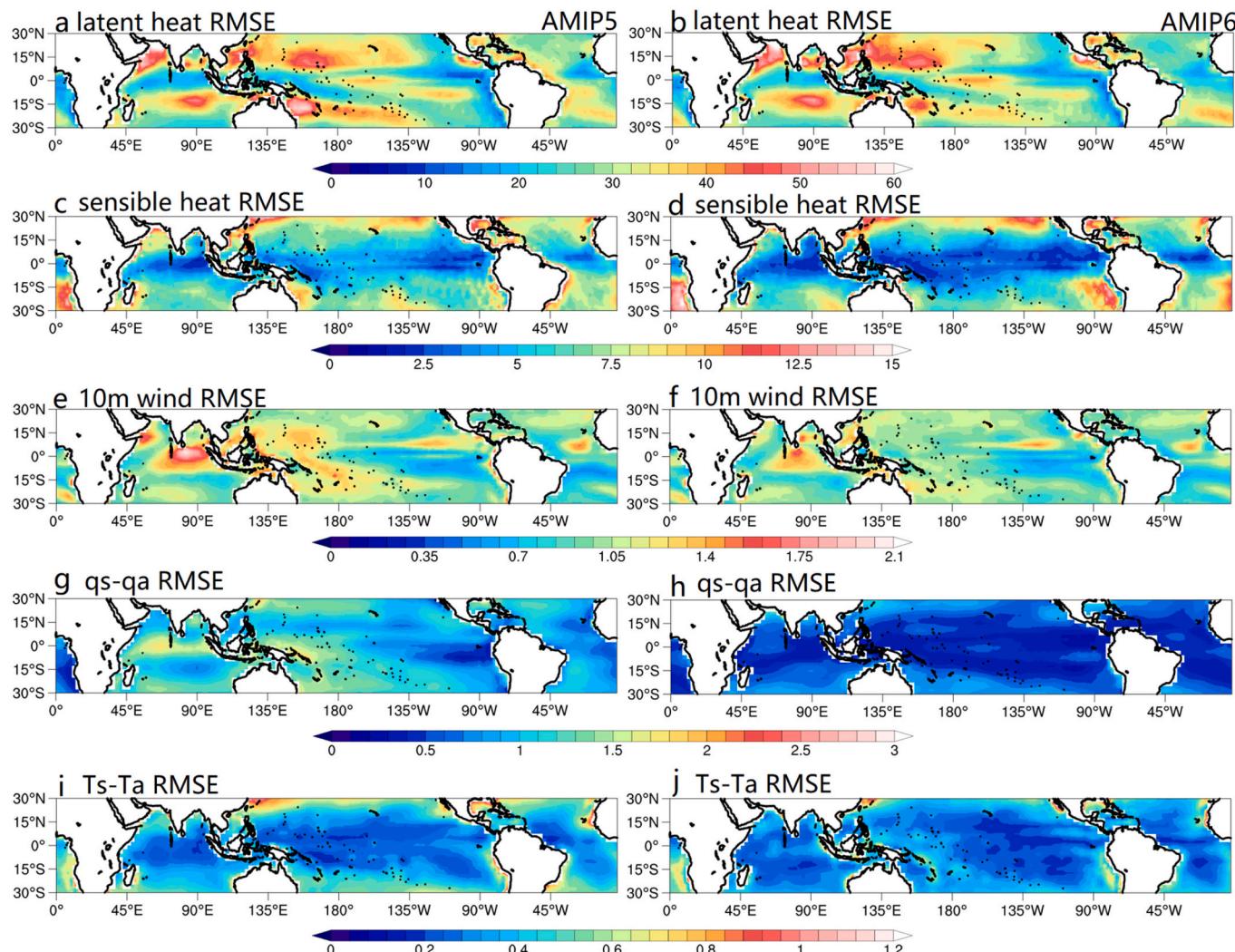


Fig. 2. The RMSE (1979–2008) of the (a-b) Q_{LH} (W m^{-2}), (c-d) Q_{SH} (W m^{-2}), (e-f) 10 m wind speed (m s^{-1}), (g-h) 2 m humidity (g kg^{-1}) and (i-j) 2 m temperature (K) from AMIP5 and AMIP6 compared to OAFlux.

Table 2

The bias (W m^{-2}), RMSE (W m^{-2}), and CC for latent heat flux (Q_{LH}) and sensible heat flux (Q_{SH}) from AMIP5 (or A5) and AMIP6 (or A6) against OAFlux. Model numbers are following Table 1. Model numbers in bold indicate models with finer horizontal resolutions (< 1.5°). All are averaged over 30°S-30°N.

Model #	Q_{LH}	Q_{SH}										
		Bias		RMSE		CC		Bias				
		A5	A6	A5	A6	A5	A6	A5	A6			
1	20	17	30	26	0.93	0.94	3	2	8	8	0.71	0.73
2	19	17	28	27	0.92	0.92	6	4	10	9	0.65	0.68
3	20	17	29	27	0.91	0.91	6	5	11	10	0.68	0.70
4	15	15	26	25	0.93	0.94	3	2	7	7	0.73	0.74
5	27	25	35	33	0.94	0.95	4	4	8	7	0.72	0.73
6	26	24	35	31	0.93	0.93	10	8	16	13	0.77	0.77
7	26	23	33	31	0.94	0.94	4	3	7	7	0.78	0.78
8	18	16	26	24	0.93	0.93	3	3	7	6	0.74	0.75
9	18	17	27	26	0.93	0.93	3	2	7	7	0.79	0.81
Ensemble (all models)	21	19	30	28	0.93	0.94	5	4	9	8	0.73	0.75
Ensemble (finer resolutions)	20	18	29	27	0.93	0.94	3	2	7	7	0.75	0.77
Ensemble (coarser resolutions)	22	20	31	29	0.92	0.93	6	5	10	9	0.71	0.73

Table 3

The bias (W m^{-2}), RMSE (W m^{-2}), and CC for latent heat flux (Q_{LH}) and sensible heat flux (Q_{SH}) from AMIP5 (or A5) and AMIP6 (or A6) against OAFlux. Model numbers are following Table 1. Model numbers in bold indicate models with finer horizontal resolutions ($< 1.5^\circ$). All are averaged over 10°S - 10°N .

Model #	Q_{LH}						Q_{SH}					
	Bias		RMSE		CC		Bias		RMSE		CC	
	A5	A6	A5	A6	A5	A6	A5	A6	A5	A6	A5	A6
1	18	16	36	25	0.95	0.95	2	1	4	3	0.81	0.83
2	17	16	26	24	0.93	0.95	5	2	7	6	0.78	0.79
3	17	15	26	25	0.92	0.94	5	4	6	7	0.75	0.76
4	14	14	24	23	0.95	0.96	2	1	4	4	0.82	0.85
5	26	24	32	32	0.96	0.97	2	2	5	4	0.80	0.83
6	25	23	33	30	0.94	0.95	9	6	13	9	0.84	0.86
7	23	21	31	29	0.94	0.96	3	2	4	3	0.85	0.88
8	17	15	24	23	0.94	0.94	3	2	4	3	0.83	0.85
9	17	15	23	22	0.94	0.95	3	1	3	3	0.89	0.91
Ensemble (all models)	19	18	28	26	0.94	0.95	4	2	6	5	0.82	0.84
Ensemble (finer resolutions)	18	17	28	25	0.94	0.95	2	1	4	3	0.84	0.87
Ensemble (coarser resolutions)	21	19	28	27	0.94	0.95	5	3	7	6	0.80	0.82

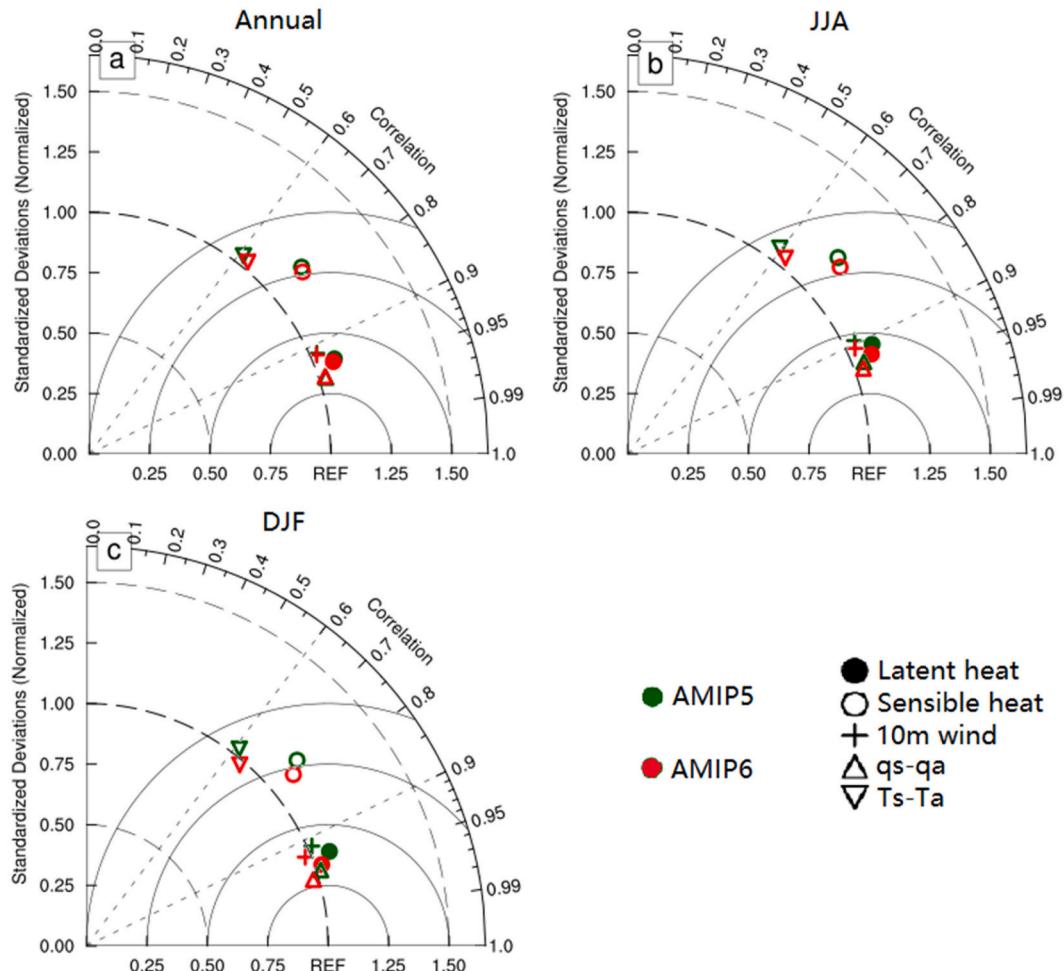


Fig. 3. Taylor diagrams describing the climatological annual mean of Q_{LH} and Q_{SH} with related variables including near-surface wind speed (U), and difference in sea surface humidity (q_s) and near-surface air specific humidity (q_a), simulated by the ensemble based on (a) nine AMIP5 models and (b) nine AMIP6 models compared to OAFlux.

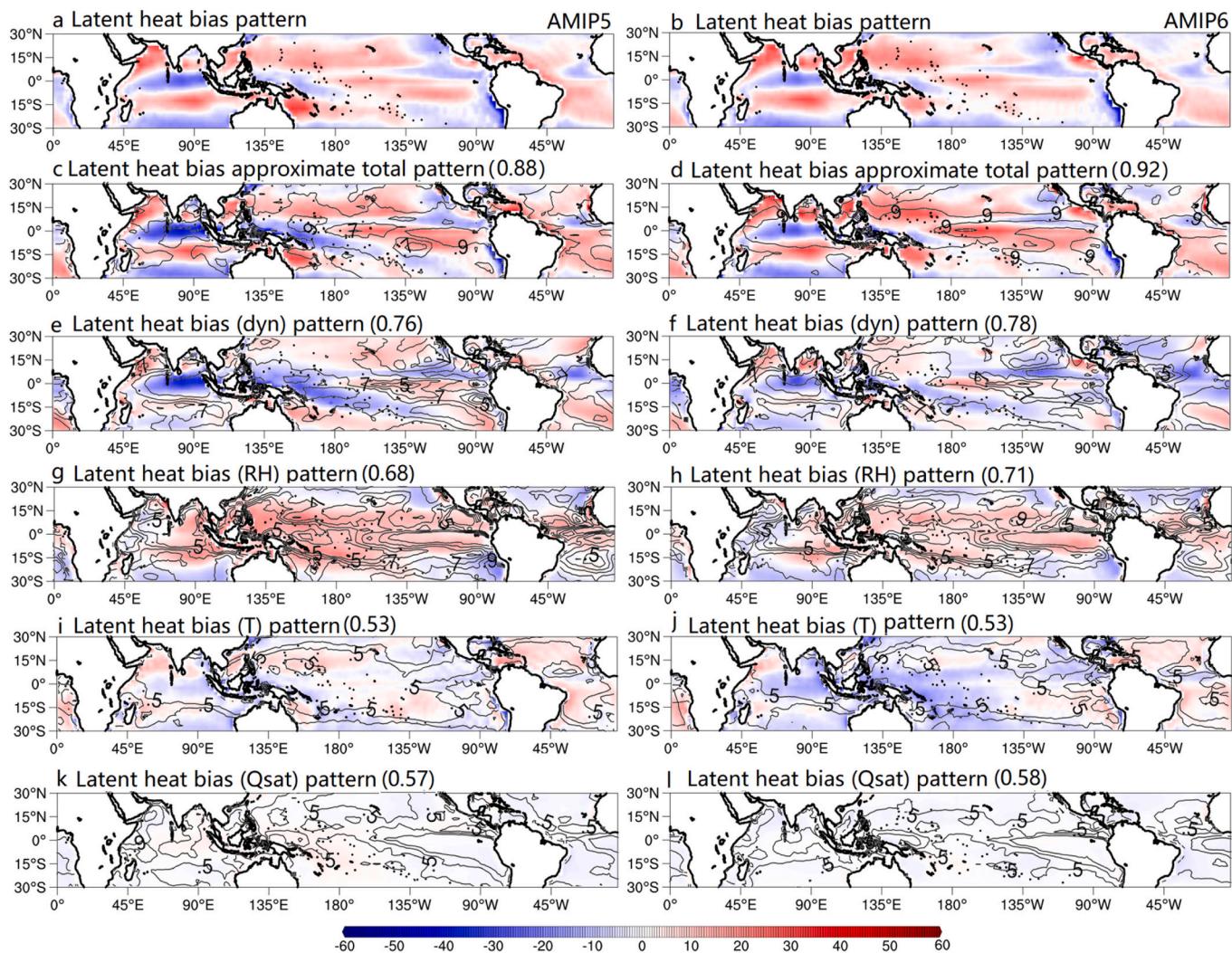


Fig. 4. (Left panels) (a) Latent heat flux composite bias pattern for the AMIP5 models and (c) total reconstructed bias pattern based on eq. 4, which is contributed by the bias pattern associated with the (e) wind speed, (g) relative humidity, (i) sea-air temperature contrast, and (k) saturation specific humidity. The contours in (c) show correlation (CC) between reconstructed total bias and actual bias, and contours in (e,g,i,k) show CC between the bias contributed by individual factors and the reconstructed total bias. The contour interval is 0.1. Numbers in bracket indicate the mean of CC in each panel. Right panels are for AMIP6 models. All calculations are over tropical oceans (30°S - 30°N) during 1979–2008. (Unit: W m^{-2}).

$$q_{\text{sat}}(T, p) = e_s(T)/p \quad (9)$$

where e_s is the saturation vapor pressure of water vapor (Pa), p is atmospheric pressure (Pa), T is air temperature (K). The e_s is given by

$$\Delta SH_{\delta T} = \rho C_p |V_a| C_H \Delta \delta T \quad (13)$$

We also calculate the “bias pattern” in any parameter Q as $(Q_{\text{AMIP}} - \overline{Q}_{\text{obs}}) - (\overline{Q}_{\text{AMIP}} - \overline{Q}_{\text{obs}})$ following Găinușă-Bogdan et al. (2018). The

$$\log(e_s(T)) = 10.80(1 - T_t/T) - 5.03\log(T/T_t) + 1.50 \times 10^{-4}(1 - 10^{(-8.30(T/T_t-1))}) + 0.43 \times 10^{-3}(10^{4.77(1-T_t/T)} - 1) + 2.79$$

where $T_t = 273.16$ K.

Similarly, the Q_{SH} and the bias coming from bulk parameters can be expressed as:

$$Q_{SH} = \rho C_p |V_a| C_H \delta T \quad (10)$$

$$\Delta Q_{SH} = \Delta SH_{\text{dyn}} \Delta HS_{\delta T} \quad (11)$$

$$\Delta SH_{\text{dyn}} = \rho C_p \Delta |V_a| C_H \delta T \quad (12)$$

removal of mean bias (shown as overbar) over the chosen study area helps to reduce uncertainty in the observed mean, in particular, for observed latent heat flux that may vary in different flux products (e.g., Bourassa et al., 2008; Tomita et al., 2010; Smith et al., 2011; Găinușă-Bogdan et al., 2015).

The method is applied to both AMIP5 and AMIP6 model output compared to OAFlux (AMIPs minus OAFlux). We primarily discuss the results for the multi-model ensemble (MME) mean, but also discuss individual models when necessary. For simplicity, MME based on AMIP5 and AMIP6 models are referred to as AMIP5 and AMIP6, unless

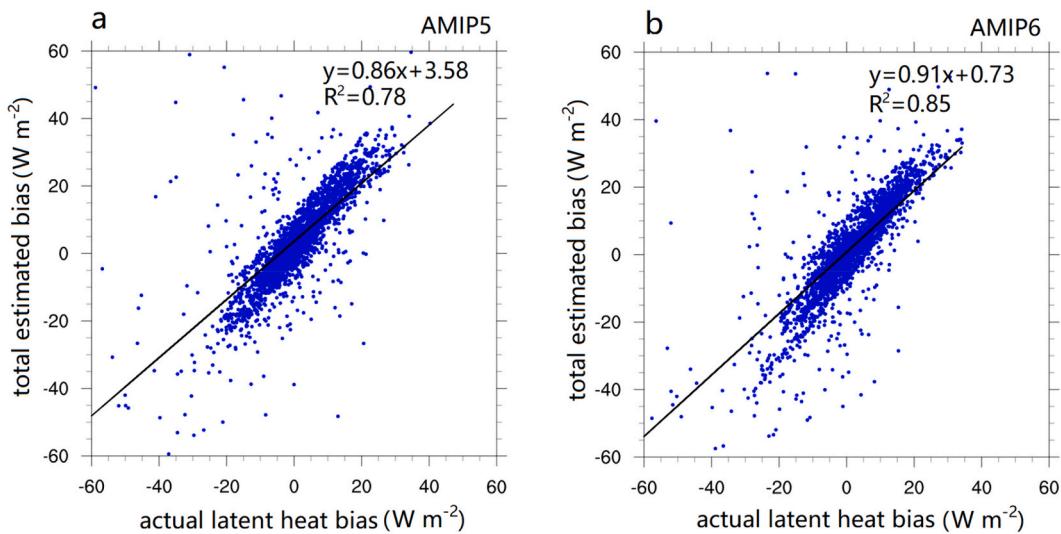


Fig. 5. Relationship between total estimated bias based on eq. (4) and actual bias in latent heat flux (W m^{-2}) for (a) AMIP5 and (b) AMIP6 using linear regression over tropical oceans (30°S - 30°N) during 1979–2008. R^2 represents square of the Pearson correlation coefficient.

Table 4

The RMSE from AMIP5 and AMIP6 model ensemble over 30°S - 30°N associated with the wind speed (Dyn), relative humidity (RH), sea-air temperature contrast (T), and saturation specific humidity (Q_{sat}). Finer (high-res $<1.5^{\circ}$) and coarser resolution (low-res $>1.5^{\circ}$) model results are also shown. (Units: W m^{-2}).

Q_{LH}						Q_{SH}					
AMIP5			AMIP6			AMIP5			AMIP6		
all	<1.5	>1.5	all	<1.5	>1.5	all	<1.5	>1.5	all	<1.5	>1.5
Dyn	12	10	13	11	10	12	1	1	1	1	1
RH	9	8	9	9	9	9	5	5	5	4	4
T	5	5	5	5	5	5	5	5	5	4	4
Q_{sat}	3	3	3	2	2	2	4	4	4	4	4

otherwise mentioned. We also use heat loss from the ocean (i.e., heat gain for the atmosphere) as positive for both latent and sensible heat flux.

4. Results

4.1. Climatological mean and seasonal cycle

The latent heat flux from OAFlux (Fig. 1a) shows off-equatorial maxima in the subtropical high where precipitation is lower than evaporation (e.g., Lau et al., 2009; Long et al., 2021). The minimum in Q_{LH} is located along the equatorial eastern and central Pacific due to the upwelling of colder water (Zhang and McPhaden, 1995). The observed sensible heat flux (Fig. 1d) also shows a pattern similar to latent heat flux (Fig. 1a). Both AMIP5 (Figs. 1b,e) and AMIP6 (Figs. 1c,f) capture the spatial patterns of Q_{LH} and Q_{SH} but overestimate both quantities. Overall, however, the AMIP6 shows a lower RMSE (Fig. 2) and higher correlation coefficient (CC, between model ensemble mean and OAFlux) (Table 2) than AMIP5. The bias (in W m^{-2}) and CC from AMIP6 in Q_{LH} (19, 0.94) and Q_{SH} (4, 0.75) are better than AMIP5 in Q_{LH} (21, 0.93) and Q_{SH} (5, 0.73) over 30°S - 30°N . For the deep tropics (10°S - 10°N), AMIP6 has bias and CC in Q_{LH} (18, 0.95) and Q_{SH} (1, 0.87) that outperform AMIP5 models (Table 3).

For bulk parameters (V_a , $q_s - q_a$, and $T_s - T_a$) that constitute the Q_{LH} and Q_{SH} , the AMIP6 is closer to OAFlux than AMIP5, especially over the tropical Indian Ocean and western Pacific Ocean (Figs. 1, g-o). In general, bulk parameters such as the equatorial 10-m wind speed (Figs. 1 g-i) and $q_s - q_a$ (Figs. 1 j-l), and off-equatorial $T_s - T_a$ (Figs. 1 m-o) are overestimated. This overestimation of 10-m wind speed and $q_s - q_a$ seen in Fig. 1 is consistent with the overestimation of Q_{LH} (Tables 2, 3; Fig. 2)

in models. For Q_{LH} (Figs. 2a,b), the RMSE value peaks in the western Arabian Sea and off-equatorial areas of the western Pacific Ocean. This pattern is more related to the 10-m wind speed (Figs. 2e,f) than $q_s - q_a$, indicating the possibility that the 10-m wind speed has a larger impact on the RMSE pattern of Q_{LH} than $q_s - q_a$. For Q_{SH} (Figs. 2c,d), the RMSE from AMIP6 is clearly lower than that from AMIP5, especially over the equatorial Oceans. This is likely because of the lower RMSE in $T_s - T_a$. The performance of the AMIP5 and AMIP6 models is summarized in the Taylor diagram (Fig. 3), where we can see that AMIP6 performed better than AMIP5, but only slightly. The causes behind the regional distribution of model bias, and its improvement in AMIP6 compared to AMIP5, are discussed next.

4.2. Causes behind the model bias

4.2.1. Latent heat flux

The bias patterns in simulated latent heat flux in AMIP5 (Fig. 4a) and AMIP6 (Fig. 4b) show overestimation in the off-equatorial Pacific Ocean, Arabian sea, Bay of Bengal and southern Indian Ocean. The bias is larger in the northwestern Pacific Ocean in AMIP6 than AMIP5, but is smaller in the southwestern Pacific. The approximate total bias (based on the method described in section 2.3) for latent heat (Figs. 4 c,d) shows high correlation (CC = 0.88 for AMIP5, and CC = 0.92 for AMIP6) and magnitude as in actual bias (Figs. 4a,b), indicating that the decomposition expressed in eq. (4) works, in agreement with earlier studies (Hourdin et al., 2015; Găinușă-Bogdan et al., 2018). Moreover, a linear regression analysis shows that the estimated bias in latent heat flux based on eq. (4) can capture the dominant variability in the bias in latent heat flux in the models (Fig. 5). This provides confidence in looking further into the bias coming from individual bulk parameters.

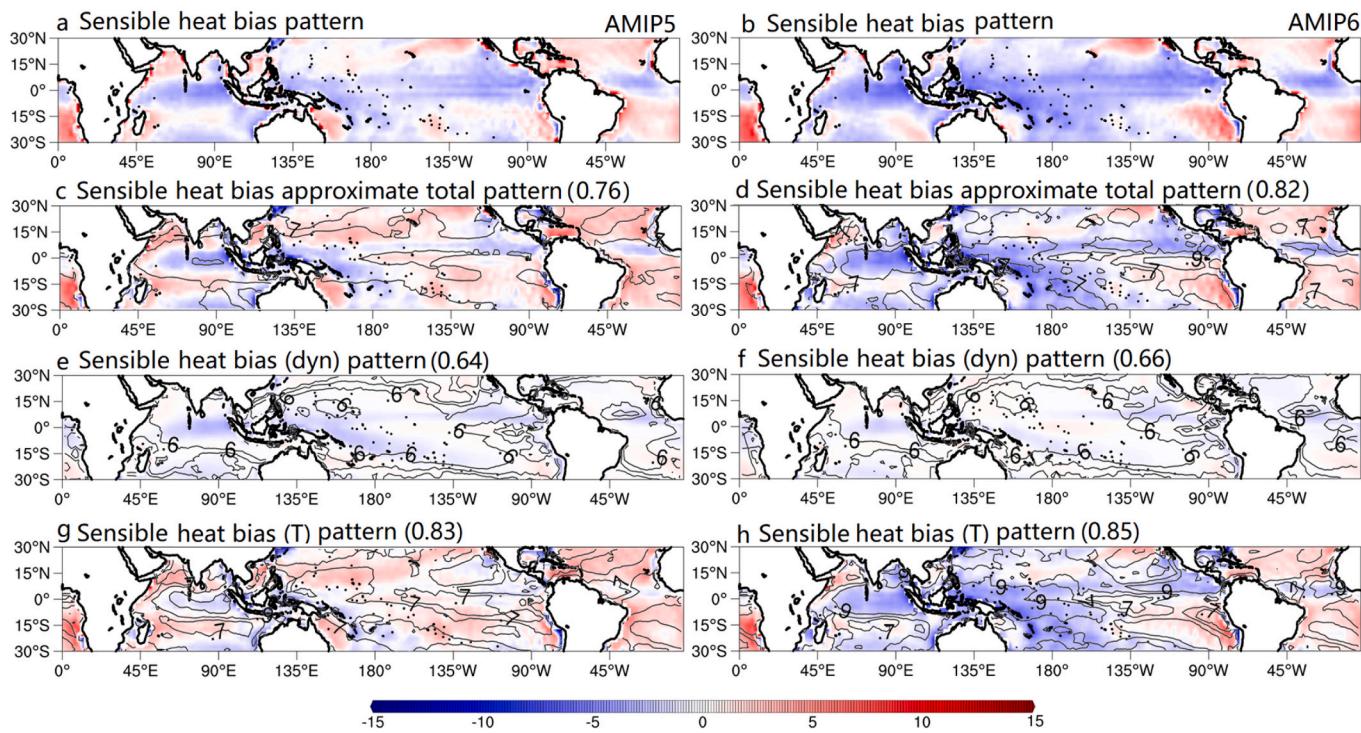


Fig. 6. (Left panels) (a) Sensible heat flux composite bias pattern for the AMIP5 models and (c) total reconstructed bias pattern based on eq. 11, which is contributed by the bias pattern associated with the (e) wind speed, and (g) sea-air temperature contrast. The contours in (c) show correlation (CC) between reconstructed total bias and actual bias, and contours in (e,g) show CC between the bias contributed by individual factors and the reconstructed total bias. The contour interval is 0.1. Numbers in bracket indicate the mean of CC in each panel. Right panels are for AMIP6 models. All calculations are over tropical oceans (30°S - 30°N) during 1979–2008. (Unit: W m^{-2}).

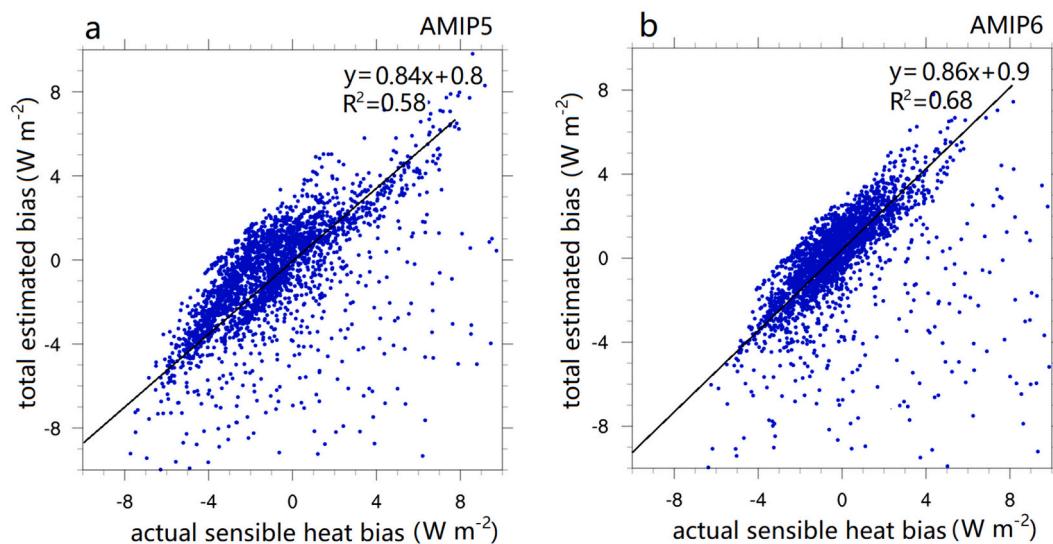


Fig. 7. Relationship between total estimated bias based on eq. (11) and actual bias in sensible heat flux (W m^{-2}) for (a) AMIP5 and (b) AMIP6 using linear regression over tropical oceans (30°S - 30°N) during 1979–2008. R^2 represents square of the Pearson correlation coefficient.

Most of the bias in latent heat seen in AMIP5 (Fig. 4c) and AMIP6 (Fig. 4d) comes from the bias in 10-m wind speed (Figs. 4e,f) and RH (Figs. 4g,h). Over the equatorial Indo-Pacific warm pool area, the pattern in Q_{LH} bias is similar to the bias pattern in 10-m wind speed in both AMIP5 and AMIP6. The contribution from the air-sea temperature contrast pattern (Figs. 4f,l) to the total bias is small in both AMIP5 and AMIP6 models compared to that coming from 10-m wind speed and RH (Table 4). The bias characteristics in Fig. 4 are consistent with that in Găinușă-Bogdan et al. (2018), who confirmed that bias from 10-m wind

speed and RH contributes most of the total bias in Q_{LH} for AMIP models. Moreover, the correlation between the approximated total bias pattern (Figs. 4c,d) and that coming from the 10-m wind speed (Figs. 4e,f) and RH (Figs. 4g,h) is higher for both AMIP5 and AMIP6 models (Fig. 4). AMIP6 shows better performance than AMIP5 with lower RMSE in Q_{LH} (Table 4). This improvement of the total bias in AMIP6 compared to AMIP5 in simulating Q_{LH} was driven by an improvement in 10-m wind speed and saturation humidity, whereas there was minor improvement from RH and T. The dependence of these results on model horizontal

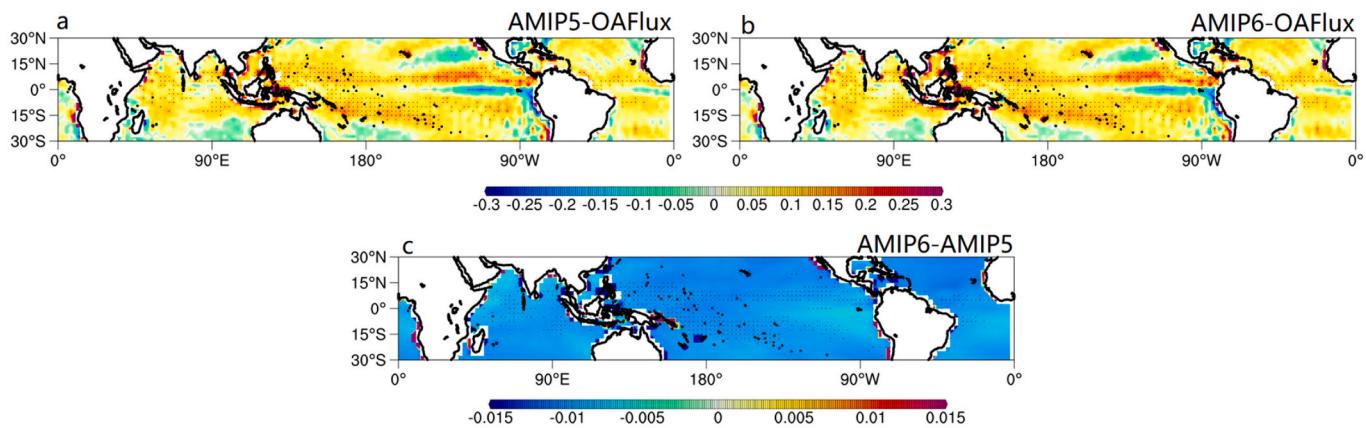


Fig. 8. Sea surface temperature for (a) AMIP5 minus OAFlux, (b) AMIP6 minus OAFlux, and (c) AMIP6 minus AMIP5 during 1979 to 2008. Dotted area indicates where the differences are significant at the 95% level based on a two-tailed Student's *t*-test. (Unit: K).

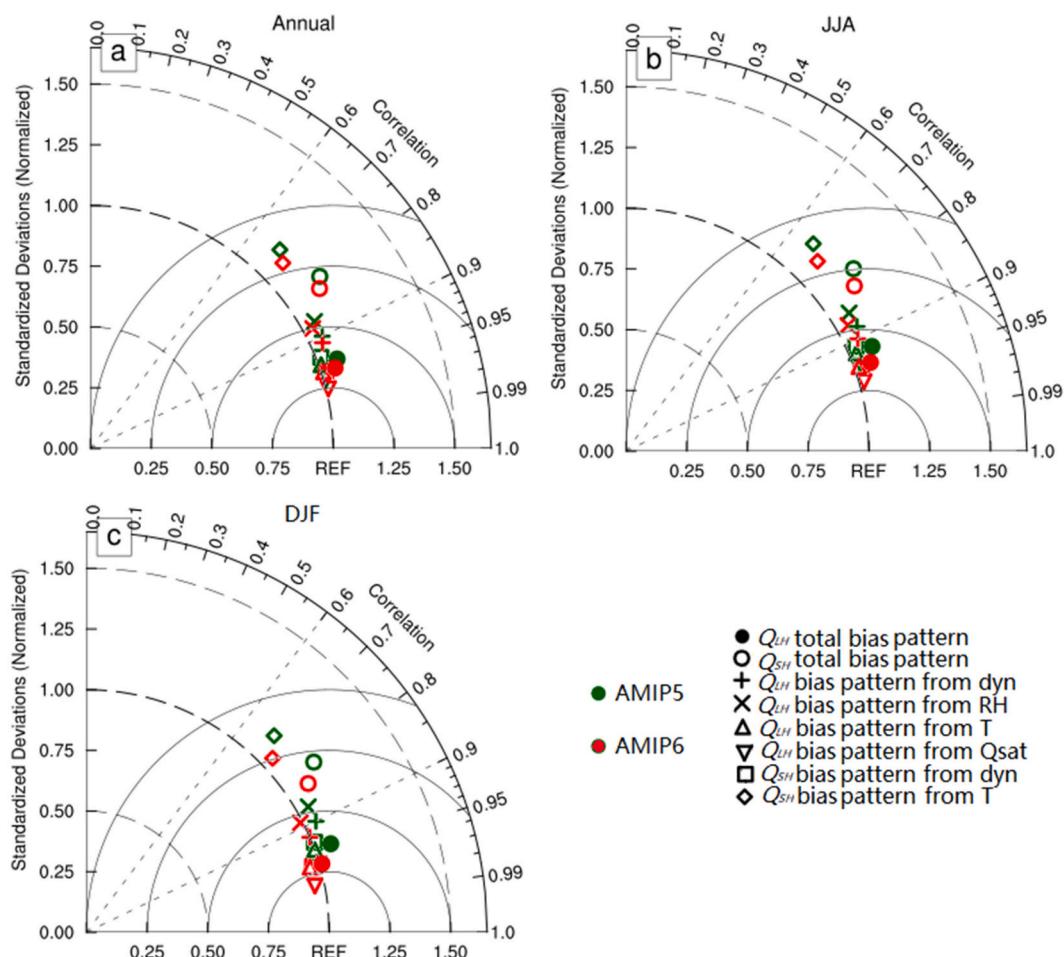


Fig. 9. Taylor diagrams describing the (a) annual mean (b) northern summer (JJA), and (c) northern winter (DJF) bias pattern of Q_{LH} and Q_{SH} , and the contribution of bias coming from the bias in the wind speed (dyn), relative humidity (RH), sea-air temperature contrast (T), and saturation humidity (Q_{sat}).

resolution is discussed in section 3.3.

4.2.2. Sensible heat flux

The bias pattern in sensible heat flux (Fig. 6, top) is also well approximated using eq. (11) in section 2.3 (Figs. 6c,d) (averaged CC = 0.76 for AMIP5, and CC = 0.82 for AMIP6). Similar to latent heat flux, a linear regression analysis for sensible heat flux shows that the estimated bias in sensible heat flux based on eq. (11) can capture the dominant

variability in the bias in sensible heat flux in the models (Fig. 7). The somewhat larger mismatch shown in Fig. 7 in the presence of higher positive values in actual bias and negative values in estimated bias comes from the coastal areas due to land-sea mask issues in the relatively coarse-resolution AMIP/OAFlux data. The underestimation in sensible heat flux along the equator, and overestimation in the off-equatorial oceans, are captured well in both AMIP5 and AMIP6. Air-sea temperature contrast (Figs. 6g,h) contributes most (with RMSE 5 and 4 W m⁻²

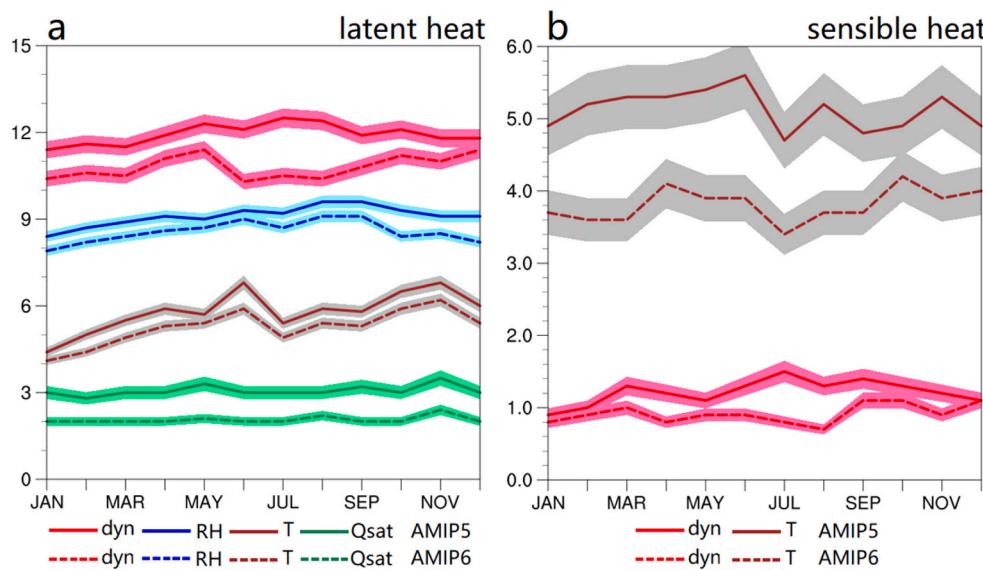


Fig. 10. (left) The annual cycle of total reconstructed latent heat flux RMSE that owes to bias from the wind speed (dyn), relative humidity (RH), the sea-air temperature contrast (T), and saturation humidity (Q_{sat}). The solid lines are for AMIP5, and the dotted lines are for AMIP6. The shaded area around the ensemble means shows the 95% confidence limits based on a two-tailed Student's t-test. The right panel is for sensible heat flux. Note that the y-axis limits are different across the two panels. (Unit: $W m^{-2}$).

Table 5

Annual and seasonal mean bias and RMSE in Q_{LH} and Q_{SH} for AMIP5 during 1997–2008. The values in parentheses are for AMIP6. Comparison with the OAFlux data was made over $30^{\circ}S\text{--}0^{\circ}$ for the Southern Hemisphere (SH) and $0^{\circ}\text{--}30^{\circ}N$ for the Northern Hemisphere (NH). Note that we have put Q_{LH} and Q_{SH} as positive numbers, which represent heat gain to the atmosphere or heat loss from the ocean. Units: $W m^{-2}$.

	DJF		MAM		JJA		SON		Annual	
	Bias	RMSE								
Northern Hemisphere										
Q_{LH}	24(23)	33(30)	22(21)	30(29)	17(16)	29(28)	20(17)	29(28)	21(19)	30(28)
Q_{SH}	4(3)	8(7)	6(5)	10(8)	7(5)	12(10)	5(4)	7(7)	5(4)	9(8)
Southern Hemisphere										
Q_{LH}	19(17)	27(26)	20(17)	28(28)	22(21)	31(30)	21(20)	30(28)	21(19)	29(28)
Q_{SH}	6(5)	11(10)	6(4)	10(8)	5(3)	8(7)	4(3)	8(7)	5(4)	9(8)

Table 6

Annual mean of Q_{LH} and Q_{SH} from AMIP5 and AMIP6 model ensemble and their bias compared to the OAFlux (1979–2008) for finer resolution (high-res $< 1.5^{\circ}$) and coarser resolution (low-res $> 1.5^{\circ}$) models. Comparison with the OAFlux was made over $30^{\circ}S\text{--}30^{\circ}N$ and $10^{\circ}S\text{--}10^{\circ}N$ (parentheses). Units: $W m^{-2}$.

Obs	AMIP5	AMIP6	AMIP5 Bias	AMIP6 Bias
Comparison with finer resolution models				
Q_{LH}	116 (107)	136 (125)	134 (124)	20 (18)
Q_{SH}	7 (5)	10 (7)	9 (6)	3 (2)
Comparison with coarser resolution models				
Q_{LH}	116 (107)	138 (128)	136 (126)	22 (21)
Q_{SH}	7 (5)	13 (10)	12 (8)	6 (5)

for AMIP5 and AMIP6 in Table 4) to the Q_{SH} bias pattern, whereas the 10-m wind speed (Figs. 6e,f) contributes $1 W m^{-2}$ in both models. The bias pattern is also dominated by differences in air-sea temperature (Fig. 6g,h), and is slightly more prominent in AMIP6 (CC = 0.85) compared to that in AMIP5 (CC = 0.83).

To further understand the role of air-sea temperature difference, we show the spatial structure of SST difference between AMIP5, AMIP6 and OAFlux from 1979 to 2008 (Fig. 8). It is not a surprise that the difference in SST between AMIP5 and OAFlux (Fig. 8a) and between AMIP6 and OAFlux (Fig. 8b) is small. However, the SST from both models is generally higher than that from OAFlux, and is expected to produce higher latent and sensible heat loss than OAFlux. The SST is also very similar between AMIP5 and AMIP6 models (Fig. 8c). For example, the

difference between AMIP5 and AMIP6 is less than 0.015 K over most areas in the tropical oceans. However, the SST is cooler in AMIP6 than in AMIP5 models in most part of the tropical oceans. This cooler SST is likely responsible for some reduction in sensible and latent heat flux in AMIP6 than in AMIP5, thereby reducing the bias in both fluxes in AMIP6 compared to AMIP5 (Tables 2 and 3). Therefore, the improvement in simulated fluxes in AMIP6 is, at least partly due to cooler SST.

The contribution from bulk parameters for both models towards annual and seasonal bias is summarized in Fig. 9. Overall, AMIP6 performs better than AMIP5 as the annual (Fig. 9a) and seasonal means (Fig. 9b,c) of heat fluxes and bulk variables are closer to observation. The humidity bias contributes more to bias in Q_{LH} and the air-sea temperature contrast bias contributes more to bias in Q_{SH} . These results are consistent with Zhou et al. (2020).

4.2.3. Annual cycle and hemispheric differences

Fig. 10 shows the annual cycle of RMSE in Q_{LH} and Q_{SH} that owes to bias from different bulk variables. We found that the RMSE in both flux components in AMIP6 is smaller than that in AMIP5 in all variables (V_a , RH, T and Q_{sat}) in every month. The biggest difference between AMIP5 and AMIP6 in Fig. 10a is the RMSE from the wind speed during June to September, part of which is likely coming from a better simulation of the monsoon winds (Rajendran et al., 2021) in AMIP6. The RMSE from relative humidity and air-sea temperature difference for both AMIP5 and AMIP6 models is small. Similarly, for Q_{SH} (Fig. 10b), the biggest difference between AMIP5 and AMIP6 in RMSE comes from the air-sea temperature contrast. These results are summarized in Table 5. The mean bias and RMSE of Q_{LH} and Q_{SH} from AMIP6 are always smaller than that from AMIP5 in both hemispheres (not shown).

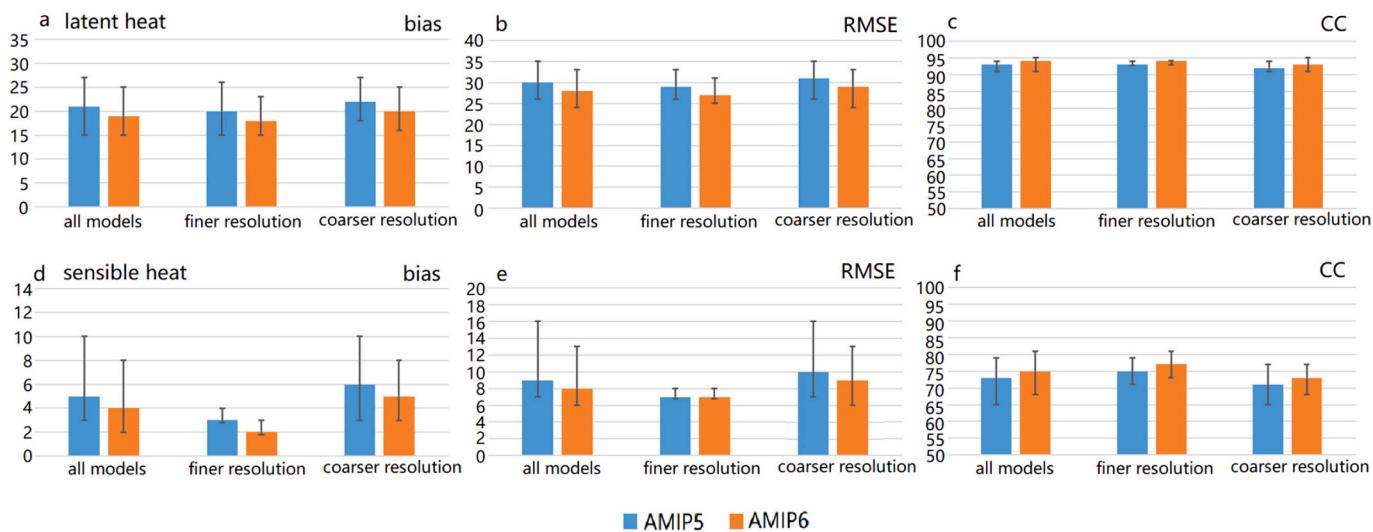


Fig. 11. (top) (a) Bias (W m^{-2}), (b) RMSE (W m^{-2}), and (c) CC for Q_{LH} in AMIP5 (blue) and AMIP6 (orange) models with finer horizontal resolution (<1.5° grid spacing) and coarser horizontal resolution (>1.5° grid spacing). (bottom) (d-f) are for Q_{SH} . The top and bottom of the black bars show the highest and lowest values from the models. Note that the y-axis limits are different across panels. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.3. Dependence on model resolution and other factors

To better understand the role of horizontal resolution on the simulated latent and sensible heat flux, we split the AMIP5 and AMIP6 models into two groups. One group contains models with the horizontal resolution finer than 1.5° (high-res), and the other group contains models with the horizontal resolution coarser than 1.5° (low-res). The results (Table 6) indicate that the biases in Q_{LH} and Q_{SH} from high-res models from AMIP5 are 20 and 3 W m^{-2} , respectively. Biases in Q_{LH} and Q_{SH} from low-res models are 22 and 6 W m^{-2} . The CC between the high-res models and OAFlux is higher than the low-res models (Tables 2–3). The results are similar for AMIP6. Models with higher resolution from AMIP6 perform better than models from AMIP5 for both Q_{LH} and Q_{SH} (Fig. 11) over most of the tropical oceans. It is clear that models with finer horizontal resolutions have lower bias and RMSE (Figs. 11a,b,d,e) and CC (Figs. 11c,f). These results are consistent with past studies (e.g., Demory et al., 2014; Vannière et al., 2018), where an improvement in the simulated surface heat flux was found due to an increased resolution of the models.

Apart from the horizontal resolution of the models, the number of vertical layers, particularly the number of vertical layers in the planetary boundary layer, has been found to influence boundary layer mixing and model convection (e.g., Aligo et al., 2009). Unfortunately, the AMIP5 and AMIP6 models provide data only at 17 and 19 specified levels, respectively, and do not provide information on the number of vertical layers in the planetary boundary layer or bottom 1000 m. Therefore, the possibility that some of the better-performing models may have higher number of vertical layers in the planetary boundary layer cannot be ruled out.

There are other factors apart from the model resolution that may also influence our results. For example, the estimation of fluxes in OAFlux and models may have some differences. First, the monthly values of the fluxes in OAFlux are calculated from their daily flux values (OAFlux doesn't have sub-daily data). The monthly AMIP fluxes, on the other hand, are produced based on sub-daily flux values from model output. So, there are some differences in the way monthly heat fluxes are produced in the AMIP and OAFlux. However, this does not affect our conclusion since a small change in OAFlux will not change our results regarding the relative performance of AMIP5 and AMIP6 compared to OAFlux. Second, the fluxes in OAFlux and AMIP datasets are based on instantaneous values of bulk variables unlike some other sophisticated

models such as the Weather Research and Forecasting (WRF) model, where fluxes are calculated at each model time-step and the time-integrated values of the fluxes can be taken as the model output (e.g., Brownlee et al., 2017; Ray et al., 2021). Therefore, our model-data comparison, although consistent, may also introduce some uncertainties. Third, the results related to mean bias will not change if we use daily flux values instead of monthly flux values since monthly values are simply based on monthly average of daily values. There is, however, one advantage in using monthly OAFlux data compared to daily data, since the biases in the monthly OAFlux data are much smaller than the daily data (Yu et al., 2008; Pinker et al., 2014). Therefore, the biases in OAFlux are expected to be even smaller for seasonal and annual means.

5. Conclusion

This study evaluates the simulated surface latent (Q_{LH}) and sensible (Q_{SH}) heat fluxes over the tropical oceans over 30°S – 30°N during a 30 yr period (1979–2008) from nine contemporary AGCMs participating in the AMIP5 and AMIP6. It is quite rare to find studies related to surface flux evaluation in atmospheric models, even though understanding the performance of atmospheric models is critical to understanding the performance of their coupled counterparts. The AMIP models used here had different resolutions, but they had the same resolutions in two versions of the AMIP simulations (i.e., AMIP5 and AMIP6). This allows us to quantify and understand any improvement in the simulated Q_{LH} and Q_{SH} in AMIP6 compared to AMIP5 that is independent of model resolution. The main conclusions can be summarized as follows:

- (1) Both AMIP5 and AMIP6 model ensembles overestimate Q_{LH} and Q_{SH} (Tables 2–3), especially over the off-equatorial west Pacific and the southeastern Indian Ocean, the east coast of Australia, and the north Indian Ocean (Fig. 2). The performance of AMIP6 was better than AMIP5. However, systematic spatial biases remained, which led to only incremental improvements in the simulated fluxes. For example, the RMSE and mean bias in Q_{LH} are 30 and 21 W m^{-2} in AMIP5 and 28 and 19 W m^{-2} in AMIP6. This represents a reduction in Q_{LH} bias by 10% and RMSE by 7%. For Q_{SH} , the improvements represent a reduction in Q_{SH} bias by 25% and RMSE by 13%.
- (2) Each AMIP6 model performed better than their AMIP5 counterparts with the same horizontal resolutions (Tables 2–3). This is

rather remarkable, and is partly due to better boundary conditions (e.g., sea surface temperature as shown in Fig. 8) (Eyring et al., 2016) and model physics in AMIP6 than AMIP5. Earlier studies related to flux evaluation (e.g., Cao et al., 2015; Bentamy et al., 2017; Găinușă-Bogdan et al., 2018) could not conclude whether the improvement was due to increased resolution, model physics or model boundary conditions.

Although we compared the models with same horizontal resolutions, four AMIP6 models have higher vertical resolutions compared with AMIP5 (Table 1). These four AMIP6 models with higher vertical resolution performed similarly to the rest of the AMIP6 models, indicating that the increased vertical resolution was not instrumental in better performance of the AMIP6 models.

- (3) To obtain insights into the mechanisms responsible for the improvement in AMIP6 compared to AMIP5, we examined the contribution of bias in Q_{LH} and Q_{SH} coming from 10-m wind speed, RH, air-sea temperature contrast, and saturation specific humidity. Most Q_{LH} bias is due to 10-m wind speed and RH (Table 4, Fig. 4), and most Q_{SH} bias is due to air-sea temperature difference (Table 4, Fig. 6). These results are consistent with Găinușă-Bogdan et al. (2018). The improvement in Q_{LH} in AMIP6 is due to a better simulation of the 10-m wind speed and saturation specific humidity (Table 4, Fig. 4). Likewise, the improvement in Q_{SH} in AMIP6 is due to a better simulation of the air-sea temperature difference (Table 4, Fig. 6).
- (4) Models with finer horizontal resolution generally performed better than those with coarser resolution (Table 6, Fig. 11). This result is consistent with previous studies (e.g., Demory et al., 2014; Vannière et al., 2018).

Overall, the AMIP6 models showed a small but systematic improvement in simulating the surface latent and sensible heat fluxes compared to AMIP5 models of same resolutions, although systematic spatial biases remained. It would be interesting to explore the performance of coupled models with same resolutions from CMIP5 and CMIP6. Such an initiative will be undertaken in the future.

CRediT authorship contribution statement

Xin Zhou: Formal analysis, Methodology, Writing – original draft. **Pallav Ray:** Conceptualization, Writing – original draft, Funding acquisition. **Bradford S. Barrett:** Writing – review & editing. **Pang-Chi Hsu:** Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was partially supported by a grant from the Office of Naval Research (N00014-16-01-3091) to PR. The authors are grateful for the surface heat flux data provided by WHOI OAFlux project (<http://oaflux.whoi.edu>). We thank Charlotte Demott for providing useful feedback on this work. We thank the climate modeling groups (listed in Table 1 of this paper) for producing their model output and making it available.

References

- Aligo, E.A., Gallus Jr., W.A., Segal, M., 2009. On the impact of WRF model vertical grid resolution on midwest summer rainfall forecasts. *Weather Forecast.* 24 (2), 575–594. <https://doi.org/10.1175/2008WAF2007101.1>.
- Bauer, S.E., Tsigaridis, K., Faluvegi, G., Kelley, M., Lo, K.K., Miller, R.L., Nazarenko, L., Schmidt, G.A., Wu, J., 2020. Historical (1850–2014) aerosol evolution and role on climate forcing using the GISS ModelE2.1 contribution to CMIP6. *J. Adv. Model Earth Syst.* 12 (8) <https://doi.org/10.1029/2019MS001978> e2019MS001978.
- Bentamy, A., Piollé, J.F., Grouazel, A., Danielson, R., Gulev, S., Paul, F., Azelmat, H., Mathieu, P.P., von Schuckmann, K., Sathyendranath, S., Evers-King, H., Esau, I., Johannessen, J.A., Clayson, C.A., Pinker, R.T., Grodsky, S.A., Bourassa, M., Smith, S. R., Haines, K., Valdivieso, M., Merchant, C.J., Chapron, B., Anderson, A., Hollmann, R., Josey, S.A., 2017. Review and assessment of latent and sensible heat flux accuracy over the global oceans. *Remote Sens. Environ.* 201, 196–218. <https://doi.org/10.1016/j.rse.2017.08.016>.
- Bourassa, M.A., Hughes, P.J., Smith, S.R., 2008. Surface turbulent flux product comparison. *Flux News.* 5, 22–24.
- Brownlee, J., Ray, P., Tewari, M., Tan, H., 2017. Relative role of turbulent and radiative flux on the near-surface temperature in a single layer urban canopy model over Houston. *J. Appl. Meteorol. Climatol.* 56, 2173–2187. <https://doi.org/10.1175/JAMC-D-17-0088.1>.
- Cao, N., Ren, B., Zheng, J.Q., 2015. Evaluation of CMIP5 climate models in simulating 1979–2005 oceanic latent heat flux over the Pacific. *J. Adv. Atmos. Sci.* 32 (12), 1603–1616. <https://doi.org/10.1007/s00376-015-5016-8>.
- Danabasoglu, G., Lamarque, J.F., Bacmeister, J., Bailey, D.A., DuVivier, A.K., Edwards, J., Emmons, L.K., Falstoe, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M.M., Large, W.G., Lauritzen, P.H., Lawrence, D.M., Lenaerts, J.T.M., Lindsay, K., Lipscomb, W.H., Mills, M.J., Neale, R., Oleson, K.W., Otto-Bliesner, B., Phillips, A.S., Sacks, W., Tilmes, S., van Kampenhoult, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J.E., Kinnison, D., Kushner, P. J., Larson, V.E., Long, M.C., Mickelson, S., Moore, J.K., Nienhause, E., Polvani, L., Rasch, P.J., Strand, W.G., 2020. The Community Earth System Model Version 2 (CESM2). *J. Adv. Model Earth Syst.* 12 <https://doi.org/10.1029/2019MS001916> e2019MS001916.
- de Szoeke, S.P., Edson, J.B., Marion, J.R., Fairall, C.W., Bariteau, L., 2015. The MJO and air-sea interaction in TOGA COARE and DYNAMO. *J. Clim.* 28, 597–622. <https://doi.org/10.1175/JCLI-D-14-00477.1>.
- Demory, M.E., Vidale, P.L., Roberts, M.J., Berrisford, P., Strachan, J., Schiemann, R., Mizielski, M.S., 2014. The role of horizontal resolution in simulating drivers of the global hydrological cycle. *Clim. Dyn.* 42 (7–8), 2201–2225. <https://doi.org/10.1007/s00382-013-1924-4>.
- DeMott, C.A., Stan, C., Randall, D.A., Branson, M.D., 2014. Intraseasonal variability in coupled GCMs: the roles of ocean feedbacks and model physics. *J. Clim.* 27 (13), 4970–4995. <https://doi.org/10.1175/jcli-d-13-00760.1>.
- DeMott, C.A., Klingaman, N.P., Woolnough, S.J., 2015. Atmosphere-ocean coupled processes in the Madden-Julian oscillation. *Rev. Geophys.* 53 <https://doi.org/10.1002/2014RG000478>.
- Eyring, V., Bony, S., Meehl, G.A., Senior, C.A., Stevens, B., Stouffer, R.J., Taylor, K.E., 2016. Overview of the coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.* 9, 1937–1958. <https://doi.org/10.5194/gmd-9-1937-2016>.
- Feng, L., Smith, S.J., Braun, C., Crippa, M., Gidden, M.J., Hoesly, R., Klimont, Z., Van Marle, M., Van Den Berg, M., Van Der Werf, G.R., 2020. The generation of gridded emissions data for CMIP6. *Geosci. Model Dev.* 13 (2), 461–482. <https://doi.org/10.5194/gmd-13-461-2020>.
- Fiorino, M., . AMIP II Sea Surface Temperature and Sea Ice Concentration Observations. https://pcmdi.llnl.gov/mips/amip/AMIP2EXPDSN/BCS_OBS/amip2_bcs.html.
- Găinușă-Bogdan, A., Braconnot, P., Servonnat, J., 2015. Using an ensemble data set of turbulent air-sea fluxes to evaluate the ipsl climate model in tropical regions. *J. Geophys. Res. Atmos.* 120 (10), 4483–4505. <https://doi.org/10.1002/2014JD022985>.
- Găinușă-Bogdan, A., Frédéric, H., Traore, A.K., Braconnot, P., 2018. Omens of coupled model biases in the CMIP5 AMIP simulations. *Clim. Dyn.* 51, 2927–2941. <https://doi.org/10.1007/s00382-017-4057-3>.
- Hagos, S., Foltz, G.R., Zhang, C., Thompson, E., Seo, H., Chen, S., Capotondi, A., Reed, K. A., DeMott, C., Protat, A., 2021. Atmospheric convection and air-sea interactions over the tropical oceans: Scientific progress, challenges, and opportunities. *Bull. Am. Meteorol. Soc.* 101 (3), E253–E258. <https://doi.org/10.1175/BAMS-D-19-0261.1>.
- Hourdin, F., Găinușă-Bogdan, A., Braconnot, P., Dufresne, J.L., Traore, A., Rio, C., 2015. Air moisture control on ocean surface temperature, hidden key to the warm bias enigma. *Geophys. Res. Lett.* 42 (24), 10885–10893. <https://doi.org/10.1002/2015GL066764>.
- Hurrell, J., Hack, J., Shea, D., Caron, J., Rosinski, J., 2008. A new sea surface temperature and sea ice boundary dataset for the community atmosphere model. *J. Clim.* 21, 5145–5153. <https://doi.org/10.1175/2008JCLI2292.1>.
- Hyder, P., Edwards, J., Allan, R.P., Hewitt, H.T., Bracegirdle, T.J., Gergory, J.M., Wood, R.A., Meijers, A.J.S., Mulcahy, J., Field, P., Furtado, K., Bodas-Salcedo, A., Williams, K.D., Copsey, D., Josey, S.A., Liu, C., Roberts, C.D., Sanchez, C., Ridley, J., Thorpe, L., Hardiman, S.C., Mayer, M., Berry, D.I., Belcher, S.E., 2018. Critical southern ocean climate model biases traced to atmospheric model cloud errors. *Nat. Commun.* 9, 3625. <https://doi.org/10.1038/s41467-018-05634-2>.
- Jiang, C.L., Cronin, M.F., Kelly, K.A., Thompson, L., 2005. Evaluation of a hybrid satellite- and NWP-based turbulent heat flux product using tropical atmosphere-ocean (TAO) buoys. *J. Geophys. Res.* 110 <https://doi.org/10.1029/2004JC002824>.
- Kumar, B.P., Vialard, J., Lengaigne, M., Murty, V.S.N., McPhaden, M.J., 2012. TropFlux: air-sea fluxes for the global tropical oceans—description and evaluation. *Clim. Dyn.* 38, 1521–1543. <https://doi.org/10.1007/s00382-011-1115-0>.
- Lau, K.M., Kim, J.H., Sud, Y., 2009. Intercomparison of hydrologic processes in AMIP GCMs. *Bull. Am. Meteorol. Soc.* 77 (10), 2209–2227. [https://doi.org/10.1175/1520-0477\(1996\)077<2209:IOHPIA>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<2209:IOHPIA>2.0.CO;2).

- Li, C., Stevens, B., Marotzke, J., 2015. Eurasian winter cooling in the warming hiatus of 1998–2012. *Geophys. Res. Lett.* 42, 8131–8139. <https://doi.org/10.1002/2015GL065327>.
- Li, J., Miao, C., Wei, W., Zhang, G., Hua, L., Chen, Y., Wang, X., 2021. Evaluation of CMIP6 global climate models for simulating land surface energy and water fluxes during 1979–2014. *J. Adv. Model Earth Syst.* 13 <https://doi.org/10.1029/2021MS002515> e2021MS002515.
- Li, G., Xie, S.P., 2012. Origins of tropical-wide SST biases in CMIP multi-model ensembles. *Geophys. Res. Lett.* 39 (L22703) <https://doi.org/10.1029/2012GL053777>.
- Long, J., Liu, C., Liu, Z., Xu, J., 2021. Impact of the Anomalous Latent Heat Flux over the Kuroshio Extension on Western north American Rainfall in Spring: Interannual Variation and Mechanism. *Front. Earth Sci.* 8, 609619 <https://doi.org/10.3389/feart.2020.609619>.
- Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittrner, M., Brokopf, R., et al., 2019. Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and its response to increasing CO₂. *J. Adv. Model Earth Syst.* 11, 998–1038. <https://doi.org/10.1029/2018MS001400>.
- Pinker, R.T., Bentamy, A., Katsaros, K.B., Ma, Y., Li, C., 2014. Estimates of net heat fluxes over the Atlantic Ocean. *J. Geophys. Res.* 119, 410–427. <https://doi.org/10.1002/2013JC009386>.
- Rahaman, H., Ravichandran, M., 2013. Evaluation of near-surface air temperature and specific humidity from hybrid global products and their impact on latent heat flux in the North Indian Ocean. *J. Geophys. Res.* 118, 1034–1047. <https://doi.org/10.1002/jgrc.20085>.
- Rajendran, K., Surendran, S., Varghese, S.J., Sathyamath, A., 2021. Simulation of Indian summer monsoon rainfall, interannual variability and teleconnections: evaluation of CMIP6 models. *Clim. Dyn.* <https://doi.org/10.1007/s00382-021-06027-w>.
- Ray, P., Tan, H., Tewari, M., Brownlee, J., Ajayamohan, R.S., Barrett, B.S., 2021. Role of advection on near-surface temperature and wind in urban-aware simulations. *J. Appl. Meteorol. Climatol.* 60 (2), 201–221. <https://doi.org/10.1175/JAMC-D-20-0068.1>.
- Reichler, T., Kim, J., 2008. How Well Do Coupled Models Simulate Today's Climate? *Bull. Am. Meteorol. Soc.* 89(3), 303–312. <https://doi.org/10.1175/BAMS-89-3-303>.
- Schwingshakel, C., Hirschi, M., Seneviratne, S.I., 2018. A theoretical approach to assess soil moisture-climate coupling across CMIP5 and GLACE-CMIP5 experiments. *Earth Syst. Dyn.* 9 (4), 1217–1234. <https://doi.org/10.5194/esd-9-1217-2018>.
- Sengupta, D., Ray, P., Bhat, G.S., 2002. Spring warming of the eastern Arabian Sea and Bay of Bengal from buoy data. *Geophys. Res. Lett.* 29 <https://doi.org/10.1029/2002GL015340>.
- Shiogama, H., Abe, M., Tatebe, H., 2019. MIROC MIROC6 model output prepared for CMIP6 ScenarioMIP. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.5741>.
- Smith, R.N.B., Gregory, D., Wilson, C., Bushell, A.C., Cusack, S., 1999. CALCULATION OF SATURATED SPECIFIC HUMIDITY AND LARGE - SCALE CLOUD. UNIFIED MODEL DOCUMENTATION PAPER 29.
- Smith, S.R., Hughes, P.J., Bourassa, M.A., 2011. A comparison of nine monthly air-sea flux products. *Int. J. Climatol.* 31 (7), 1002–1027. <https://doi.org/10.1002/joc.2225>.
- Swart, N.C., Cole, J.N.S., Kharin, V.V., Lazare, M., Scinocca, J.F., Gillett, N.P., Anstey, J., Arora, V., Christian, J.R., Hanna, S., Jiao, Y., Lee, W.G., Majda, F., Saenko, O.A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., von Salzen, K., Yang, D., Winter, B., 2019. The Canadian Earth System Model version 5 (CanESM5.0.3). *Geosci. Model Dev.* 12, 4823–4873. <https://doi.org/10.5194/gmd-12-4823-2019>.
- Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.* 106, 7183–7192. <https://doi.org/10.1029/2000JD900719>.
- Tomita, H., Kubota, M., Cronin, M.F., Iwasaki, S., Konda, M., Ichikawa, H., 2010. An assessment of surface heat fluxes from J-OFURO2 at the KEO and JKEO sites. *J. Geophys. Res.* 115 (C03), 018. <https://doi.org/10.1029/2009JC005545>.
- Vannière, B., Demory, M.E., Vidale, P.L., Schiemann, R., Roberts, M.J., Roberts, C.D., Matsueda, M., Terray, L., Koenigk, T., Senan, R., 2018. Multi-model evaluation of the sensitivity of the global energy budget and hydrological cycle to resolution. *Clim. Dyn.* 52, 6817–6846. <https://doi.org/10.1007/s00382-018-4547-y>.
- Volodin, E., Mortikov, E., Gritsun, A., Lykossov, V., Galin, V., Diansky, N., Gusev, A., Kostrykin, S., Iakovlev, N., Shestakova, A., Emelina, S., 2019. INM INM-CM4-8 model output prepared for CMIP6 CMIP piControl. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.5080>.
- Wild, M., 2020. The global energy balance as represented in CMIP6 climate models. *Clim. Dyn.* 55, 553–577. <https://doi.org/10.1007/s00382-020-05282-7>.
- Wu, T., Zhang, F., Zhang, J., Jie, W., Zhang, Y., Wu, F., Li, L., Yan, J., Liu, X., Lu, X., Tan, H., Zhang, L., Wang, J., Hu, A., 2020. Beijing climate Center Earth System Model version 1 (BCC-ESM1): model description and evaluation of aerosol simulations. *Geosci. Model Dev.* 13, 977–1005. <https://doi.org/10.5194/gmd-13-977-2020>.
- Xiang, B., Zhao, M., Held, I.M., Golaz, J.-C., 2017. Predicting the severity of spurious “double ITCZ” problem in CMIP5 coupled models from AMIP simulations. *Geophys. Res. Lett.* 44, 1520–1527. <https://doi.org/10.1002/2016GL071992>.
- Yang, J., Wang, Z.-H., Huang, H.-P., 2019. Intercomparison of the surface energy partitioning in CMIP5 simulations. *Atmosphere* 2019 (10), 602. <https://doi.org/10.3390/atmos10100602>.
- Yu, L.S., Weller, R.A., 2007. Objectively analyzed air-sea heat fluxes for the global ice-free oceans (1981–2005). *Bull. Am. Meteorol. Soc.* 88 (4), 527–539. <https://doi.org/10.1175/BAMS-88-4-527>.
- Yu, L.S., Jin, X.Z., Weller, R.A., 2008. Multidecade global flux datasets from the Objectively Analyzed air-sea Fluxes (OAFlux) project: Latent and sensible heat fluxes, ocean evaporation, and related surface meteorological variables. In: Woods Hole Oceanographic Institution, OAFlux Project technical report. OA-2008-01. Woods Hole, Massachusetts, 64 pp.
- Yukimoto, S., Koshiro, T., Kawai, H., Oshima, N., Yoshida, K., Urakawa, S., Tsujino, H., Deushi, M., Tanaka, T., Hosaka, M., Yoshimura, H., Shindo, E., Mizuta, R., Ishii, M., Obata, A., Adachi, Y., 2019. MRI MRI-ESM2.0 model output prepared for CMIP6 CMIP. Earth System Grid Federation. <https://doi.org/10.22033/ESGF/CMIP6.621>.
- Zhang, G., McPhaden, M., 1995. The relationship between sea surface temperature and latent heat flux in the Equatorial Pacific. *J. Clim.* 8 (3), 589–605. [https://doi.org/10.1175/1520-0442\(1995\)008<0589:TRBSST>2.0.CO;2](https://doi.org/10.1175/1520-0442(1995)008<0589:TRBSST>2.0.CO;2).
- Zhang, R., Wang, X., Wang, C., 2018a. On the simulations of global oceanic latent heat flux in the CMIP5 multimodel ensemble. *J. Clim.* 31 (17), 7111–7128. <https://doi.org/10.1175/JCLI-D-17-0713.1>.
- Zhang, R., Wang, X., Wang, C.Z., 2018b. On the simulations of global oceanic latent heat flux in the CMIP5 multi-model ensemble. *J. Clim.* 31 (17), 7111–7128. <https://doi.org/10.1175/JCLI-D-17-0713.1>.
- Zhou, X., Ray, P., Boykin, K., Barrett, B., Hsu, P.-C., 2019. Evaluation of surface radiative fluxes over the tropical oceans in AMIP simulations. *Atmosphere* 10 (10), 606. <https://doi.org/10.3390/atmos10100606>.
- Zhou, X., Ray, P., Barrett, B.S., Hsu, P.C., 2020. Understanding the bias in surface latent and sensible heat fluxes in contemporary AGCMs over tropical oceans. *Clim. Dyn.* 55, 2957–2978. <https://doi.org/10.1007/s00382-020-05431-y>.
- Ziehn, T., Chamberlain, M.A., Law, R.M., Lenton, A., Bodman, R.W., Dix, M., Stevens, L., Wang, Y.-P., Srbivovsky, J., 2020. The Australian earth system model: ACCESS-ESM1.5. *J. South. Hemisph. Earth Syst. Sci.* 70, 193–214. <https://doi.org/10.1071/ES19035>.