JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2019MS001988

Special Section:
Presentation and Analysis of the IPSL Climate Model Used in CMIP6

Convective Boundary Layer Control of the Sea Surface Temperature in the Tropics

Frédéric Hourdin[1], Catherine Rio[2], Arnaud Jam[1], Abdoul-Khadre Traore[1], and Ionela Musat[1]

[1]Laboratoire de Météorologie Dynamique/IPSL/Sorbone Universités/CNRS, UMR 8539, Paris, France, [2]CNRM (UMR 3589 CNRS and Meteo-France), Toulouse, France

Key Points:

- Nonlocal vertical transport by boundary layer thermals controls near-surface humidity and in turn evaporative cooling over tropical oceans
- Mass flux convective parameterization dries the near-surface air more than a diffusive approach, reinforcing surface evaporation
- The tropical SST biases in the IPSL-CM6A-LR coupled model are reduced by improving representation of boundary layer convection and tuning

Correspondence to:
F. Hourdin,
frederic.hourdin@lmd.jussieu.fr

Citation:
Hourdin, F., Rio, C., Jam, A., Traore, A.-K., & Musat, I. (2020). Convective boundary layer control of the sea surface temperature in the tropics. *Journal of Advances in Modeling Earth Systems*, *12*, e2019MS001988. https://doi.org/10.1029/2019MS001988

Received 13 DEC 2019
Accepted 25 MAR 2020
Accepted article online 6 APR 2020

Abstract Using successive versions of a global climate model, we show how convective transport to the free troposphere of the humidity evaporated at the surface or, reciprocally, entrainment of dry air from the free troposphere into the mixed layer, controls surface evaporative cooling and then sea surface temperature. This control is as important as the radiative effect of boundary layer clouds on radiation. Those aspects are shown to be improved when activating a mass flux representation of the organized structures of the convective boundary layer coupled to eddy diffusion, the so-called "thermal plume model," leading to an increased near-surface drying compared to the use of turbulent diffusion alone. Controlling detrainment by air properties from just above the boundary layer allows the thermal plume model to be valid for both cumulus and stratocumulus regimes, improving the contrast in near-surface humidity between the trade winds region and East Tropical oceans. Using pairs of stand-alone atmospheric simulations forced by sea surface temperature and of coupled atmosphere-ocean simulations, we show how the improvement of the surface fluxes that arise from this improved physics projects into an improvement of the representation of sea surface temperature patterns in the coupled model, and in particular into a reduction of the East Tropical Ocean warm bias. The work presented here led to the bias reduction in sea surface temperature in the Institute Pierre Simon Laplace coupled model, IPSL-CM6A, developed recently for the 6th phase of the Coupled Model Intercomparison Project, CMIP6.

Plain Language Summary Global numerical models used to anticipate the future of our climate under global warming still suffer from significant errors, some of which are shared among all models. Among those shared errors is the tendency to predict too warm sea surface temperature over the east part of the tropical ocean in the tropics. We show how a better representation of the vertical convective transport in the first km above sea surface improves the representation of transport of dry air from the free troposphere. The drying of the near surface increases evaporation at the surface, that in turn contributes to significantly cool the sea surface in those regions.

1. Introduction

Most coupled atmosphere-ocean global models show warm biases over the east side of Tropical Oceans (ETO, see, e.g., Richter, 2015; Zheng et al., 2011). The average biases across models contributing to the Coupled Model Intercomparison Project (Taylor et al., 2012) have not changed between CMIP3 and CMIP5 (see, e.g., Xu et al., 2014), although separated by 7 years, as confirmed by the Intergovernmental Panel on Climate Change Fifth Assessment Report (IPCC AR5) (Flato et al., 2013).

The underestimation of the shadowing effect due to a lack or underestimation of stratocumulus clouds that cover those regions all year long, has for long been identified as a possible major source of sea surface warm biases (Ma et al., 1996; Richter, 2015; Yu & Mechoso, 1999). A poor representation of the oceanic surface cooling, by advection or mixing with the colder subsurface water, is also advocated as a possible source of the warm biases (Richter, 2015; Xu et al., 2014). This misrepresentation of oceanic cooling may be due to deficiencies in oceanic models themselves or to a misrepresentation of the surface wind stress (Voldoire et al., 2019). A lack of surface evaporative cooling has been identified recently by Hourdin et al. (2015) as a third origin of the warm biases. By comparing the sea surface temperature (SST) biases simulated in a coupled atmosphere-ocean model with surface flux biases in atmospheric simulations forced by SST with the same atmospheric model, Hourdin et al. (2015) showed that the SST warm bias across coupled models is strongly

correlated with latent heat biases of models with forced SST. Indeed, the models with less evaporative cooling over the ETO when forced by SST show warmer SST over the ETO in coupled simulations.

Several reasons explain why the role of evaporative cooling was not identified in previous studies. First, evaporative cooling is often overestimated over the whole tropics in stand-alone atmospheric simulations (Richter, 2015; Xu et al., 2014), hiding the fact that this overestimation is less pronounced over the ETO. Second, when considering fluxes directly in coupled atmosphere-ocean models (de Szoeke et al., 2010; Richter, 2015; Zheng et al., 2011), the evaporative cooling over the ETO is generally overestimated as a consequence of the warm SST bias. The evaporative cooling is in fact the main restoring process for the surface energy balance, when the SST increases in response to an overestimated atmospheric downard flux at the surface. The impact of evaporative cooling on SSTs can then only be highlighted by analyzing SST biases in coupled models with respect to heat flux biases in forced-by-SST stand-alone atmospheric simulations. It is also key to analyze the results in terms of anomaly, by contrasting the ETO and the rest of the tropics. Indeed, the mean global SST in global climate models is fundamentally tuned by adjusting the global mean atmospheric fluxes explicitly or implicitly (Hourdin et al., 2017). When considering anomalies, it appears that the surface flux anomaly pattern in stand-alone atmospheric simulations projects quite directly onto patterns of SST biases in coupled simulations. Găinuşă-Bogdan et al. (2018) showed that one third of the bias in SST anomaly pattern in the tropics in the CMIP5 ensemble can be attributed to the latent heat flux anomalies in stand-alone atmospheric simulations. Hourdin et al. (2015) concentrated on the ETO by defining an ETO anomaly index (ETOA) as the difference between the ETO and rest of the tropical oceans. Considering this ETOA, the linear regression between radiative plus turbulent flux in stand-alone atmospheric simulations and SST bias in coupled models is of 25 W m^{-2} per K (Figure 1e in Hourdin et al., 2015). With 12.6 W m^{-2} per K, the evaporative cooling alone contributes as much to the SST bias as the bias in short-wave (SW) radiation (lack of shadow). As mentionned previously, if considering the same relationship but taking the latent heat flux from coupled atmosphere-ocean simulations, the relationship is in fact reversed, with a slope of -7 W m^{-2} per K (Figure 2a in Hourdin et al., 2015).

In the same study, the lack of evaporative cooling over the ETO in stand-alone atmospheric simulations was shown to be positively correlated with overestimated near-surface humidity (Figure 1f in Hourdin et al., 2015, also in terms of anomaly with respect to the full tropics). Since an excess of humidity can hardly be explained by a lack of evaporation, the excess of humidity was rather interpreted as the origin of the lack of evaporation. By decomposing surface evaporation flux in terms of relative humidity, it was shown that the model ensemble variation in relative humidity is dominating the cross-model variation of latent heat estimates over the ETO. Biases in surface wind also contribute strongly to the latent heat flux biases. This effect is more pronounced on the west side of the oceanic basins and is expressed in zonal structures rather than east-west contrasts. Another proof of the role of near surface humidity was given by Găinuşă-Bogdan et al. (2018) who modified the surface exchange coefficient over the tropical oceans to artificially reduce the impact of the relative humidity bias on evaporative cooling in region of moist biases, resulting in a reduction of the classical ETO warm biases.

Hourdin et al. (2015) suggested that the excessive near-surface humidity in the ETO region may be due in some models to an underestimated upwa rd transport of the humidity evaporated at the surface by boundary layer convection, or (which is in fact equivalent) insufficient entrainment of dry tropospheric air into the boundary layer. The purpose of the present paper is to demonstrate this last point, based on sensitivity experiments to the representation of the boundary layer in a particular climate model, with improved representation of boundary layer transport and clouds.

It is known, in particular since the pioneering work of Deardorff (1966), that convective cells or plumes that scale with the depth of the boundary layer transport heat up the gradient of potential temperature, a phenomenon that cannot be captured by traditional eddy diffusion approaches. This nonlocal transport is much more efficient than local diffusion to evacuate toward the free troposphere the moisture coming from surface evaporation or, reciprocally, to entrain dry air from the free troposphere to the surface. A now classical approach to represent such nonlocal vertical transport in the convective boundary layer consists in combining eddy diffusion with a mass flux representation of vertical transport by convective cells or rolls inherited from parameterizations of cumulus convection (Arakawa & Schubert, 1974). In this approach first suggested by Chatfield and Brost (1987) and often called EDMF for "Eddy Diffusion Mass Flux," dry and

cloudy topped boundary layers are represented in a unified way (Hourdin et al., 2002; Soares et al., 2004; Siebesma et al., 2007).

The particular EDMF approach used here, the so-called "thermal plume model" (Hourdin et al., 2002), combines a mass flux representation of a single updraft and compensating downdraft with the Yamada (1983) eddy diffusion scheme, in which the eddy diffusion specification relies on the computation of a prognostic turbulent kinetic energy. This thermal plume model was shown by Hourdin et al. (2002) to better represent the convective transport in dry convective conditions than the sophisticated eddy diffusion formulation of Yamada (1983) or the counter-gradient approach of Holtslag and Boville (1993). It was also shown how a more efficient daytime convective transport dries the surface over continental surfaces by exporting the evaporated surface water more efficiently to the dry troposphere, in better agreement with observations (Cheruy et al., 2013; Diallo et al., 2017) and produces better mixed boundary layer profiles. The extension of the thermal plume model to the representation of cumulus clouds (Rio & Hourdin, 2008; Rio et al., 2010) and its coupling with a statistical cloud scheme based on a bi-Gaussian distribution of subgrid-scale total water (Jam et al., 2013) was shown to improve the representation of cumulus clouds in the LMDZ global atmospheric model (Hourdin et al., 2013).

However, the first versions of the thermal plume model were leading to the destruction of simulated stratocumulus over the ETO, and the parameterization was thus intentionally switched off in regions of strong inversion in the 5B version of LMDZ developed for CMIP5. Jam (2012) proposed a modification of detrainment of air from the thermal plume model that enables to represent dry convection, stratocumulus, transition from stratocumulus to cumulus, and cumulus regimes in a unified way over the full tropics, producing in particular a reasonable coverage of stratocumulus clouds. The modification of the scheme and its effect on clouds are described in detail in Hourdin et al. (2019). This paper also shows how the thermal plume model provides a formal parameterization of top entrainment (from the free troposphere into the boundary layer) consistent with top entrainment resolved by LES. This modification is one of the main improvements between the versions 5B of LMDZ and 6A developed for CMIP6.

In the present study, we show how the activation of the thermal plume model over the ETO increases near-surface drying and in turn the evaporative cooling compared to the previous version. By using pairs of simulation runs with the same atmospheric model either in stand-alone atmospheric mode or coupled to an oceanic model, we show how the improvement in surface flux lead to a reduced SST bias over the ETO. This work was key in the improvement of the climatology of the Coupled Model of Intitut Pierre Simon Laplace, IPSL-CM. This paper is also thought as a reference paper to document the rationale behind this improvement and is part of a special collection of the "Journal of Advances in Modeling Earth Systems" dedicated to the "Presentation and analysis of the IPSL climate model used in CMIP6" , that uses the version 6A of LMDZ.

Section 2 presents the model and setup for numerical simulations. Section 3 presents the results of a series of sensitivity experiments to atmospheric physics in both stand-alone atmospheric simulations forced by SSTs and coupled atmosphere-ocean simulations. Results are discussed in section 4.

2. Model and Methodology

The results presented in this paper were obtained with three successive versions of the LMDZ atmospheric general circulation model associated with the successive phases 5 and 6 of CMIP. Two configurations were developed for CMIP5. The first version called 5A (Hourdin et al., 2013) was based on a rather old and robust version of the physical parameterizations hereafter called "Standard Physics." A "New Physics" version (called 5B) was also developed to contribute to CMIP5 (Hourdin et al., 2013). It differs from the "Standard Physics," by the representation of the boundary layer presented below, and modification of the Emanuel (1991) deep convection scheme, which is coupled to a parameterization of the cold pools created below convective towers by the evaporation of precipitation (Grandpeix & Lafore, 2010; Grandpeix et al., 2010; Rio et al., 2009). For CMIP6, the "New Physics" version was improved in several aspects, using a new radiative scheme, introducing a stochastic triggering of deep convection (Rochetin et al., 2014) and, more importantly for the present paper, activating the thermal plume model in regions of strong inversion (Hourdin et al., 2019). This model configuration is described in a companion paper (Hourdin et al., 2020) in the same special collection. Note that all the simulations are coupled with the Orchidee land surface model (D'Orgeval et al., 2008).

2.1. Boundary Layer Parameterizations in LMDZ

Here we describe the parameterization of boundary layer transport, and its evolution from version 5A to 6A.

The subgrid scale turbulent transport of a scalar quantity q reads

$$\frac{\partial \rho q}{\partial t} = -\frac{\partial \rho \overline{w'q'}}{\partial z} \tag{1}$$

(ρ being the air density, q' and w' the turbulent fluctuations of q and w, and z and t the vertical and time coordinates).

In the version 5A of LMDZ, the vertical turbulent flux is given by

$$\rho \overline{w'q'} = -\rho K_z \left(\frac{\partial q}{\partial z} - \Gamma \right) \tag{2}$$

K_z is the turbulent diffusion coefficient, computed as a function of a mixing length, wind shear, and Richardson number (Laval et al., 1981). Γ is a countergradient term (set to +1 K/km), introduced on potential temperature θ to maintain an upward flux of heat even in a neutral or slightly stable atmosphere, as is almost systematically observed in the convective boundary layer.

In the version 5B, the computation of K_z is replaced by a prognostic equation for the turbulent kinetic energy, following the scheme by Mellor and Yamada (1974) adapted to the atmosphere by Yamada (1983). This parameterization has been revisited for version 6A to better represent strong inversion in very stable boundary layers (Vignon et al., 2018). In versions 5B and 6A of the model, the thermal plume model replaces both the countergradient term ($\Gamma = 0$ in equation (3)) and a dry adjustment parameterization, so that the subgrid scale turbulent transport of a scalar quantity q now reads

$$\rho \overline{w'q'} = -\rho K_z \frac{\partial q}{\partial z} + f\left(q_u - q\right) \tag{3}$$

where $f = \rho \alpha w$ is the vertical mass flux in the updraft fraction of the mesh α (equal to the downward mass flux in the subsiding fraction $1 - \alpha$ of the mesh), and q_u the concentration of tracer within the updrafts, computed from the continuity equation for quantity q:

$$\frac{\partial f q_u}{\partial z} = eq - dq_u \tag{4}$$

where e and d are the lateral entrainment and detrainement related to the mass flux by the air continuity equation (same equation with $q \equiv q_u \equiv 1$). Lateral entrainment and detrainment rates are defined as positive functions of buoyancy $B = g(\theta_{vu} - \theta_v)/\theta_v$ (where θ_v is the virtual potential temperature) and vertical velocity in the updrafts w_u (see Rio et al., 2010, for detailed formulations), so that entrainment is strong for strong positive buoyancy and detrainment is strong for large negative buoyancy. Detrainment also depends on the humidity contrast $\Delta r = r_u - r$ between the updraft and its environment.

The boundary layer cloud cover and liquid (or ice at high latitudes) water content is computed by introducing a subgrid scale probability distribution function (PDF) for saturation deficit, the average of which is computed from the large-scale state variables for total water and temperature. The cloud cover is computed as the integral of this PDF above saturation, and the cloud water content as the integral of the water content multiplied by the PDF. When thermal plumes are activated (i.e., when there is a positive heat flux at the surface), the PDF is bimodal, one mode being associated with the thermal plume and the other one with its environment (Jam et al., 2013). The PDF reduces to a single Gaussian function when thermals are not activated, or systematically in the "Standard Physics" version 5A.

2.2. Specific Treatment of Stratocumulus Clouds

The tendency of parameterized physics to underestimate stratocumulus clouds is a long-standing issue of global climate models. Several successful studies in that respect advocate for switching parameterizations off and on between very stable or more unstable conditions (Bretherton & Park, 2009; Lock, 2001). Similarly in version 5A, different formulations for the K_z coefficient were used, depending on a threshold for the strength of the boundary layer top inversion (Hourdin et al., 2006). The threshold was fixed to $-\partial\theta/\partial p$=0.02 K/Pa. Inversions larger than this threshold value were mainly encountered in the subsidence regions, over

the ETO, where the K_z coefficient was set to zero at the top of the boundary layer, avoiding spurious vertical diffusion that would have destroyed the stratocumulus deck.

The version of the thermal plume model used in the 5B version of the model was destroying stratocumulus clouds, while they were abundant when using the (Yamada, 1983) scheme alone to compute boundary layer transport. A similar threshold on the strength of the boundary layer top inversion was thus introduced, above which the thermal plume model is switched off. The threshold value of the 5B version was adjusted to 0.08 K/Pa. Among other problems discussed below, the threshold approach introduces discontinuities in the system: over a time sequence or among neighbors columns, the physical content of the model itself can be different.

Following a proposition by Jam (2012), the thermal plume model was modified from version 5B to version 6A in order to enable a unified representation of stratocumulus clouds and transition between stratocumulus and cumulus. The detrainment formulation is modified to consider that the overshooting thermal plume sees environmental air coming from a distance l above the detrainment level: $B'(z) = g\left(\theta_{v_u}(z) - \theta_v(z+l)\right)/\theta_v(z)$ with $l = (1+D)z$, D being a free model parameter, with typical values of 0.05 to 0.15. The value of the D parameter controls the boundary layer top entrainment, and in turn the rate of deepening of the boundary layer in the transition cases, faster for smaller values of D. As explained by Hourdin et al. (2019), this modification makes the detrainment sensitive to the conditions of the free troposphere before the plume reaches cloud top. It directly impacts cloud top entrainment of dry air that was coming from too far above the inversion layer via the compensating subsidence in the 5B version, leading to the lack of simulated stratocumulus. The modification allowed to satisfactorily simulate test cases of stratocumulus and transition from stratocumulus to cumulus in single-column mode.

2.3. Model Configurations and Metrics

The present study is based on pairs of stand-alone atmospheric simulations with imposed SST and coupled ocean-atmosphere simulations with the IPSL coupled model (Boucher et al., 2020).

The stand-alone atmospheric simulations are forced at the surface by the mean climatology of the seasonal cycle of the "amip" SSTs (Taylor et al., 2000). They are evaluated in terms of radiative and turbulent fluxes, near-surface relative humidity, low-level cloud cover, and temperature contrast between the surface and the air above. The second edition of the CERES-EBAF L3b product is used as a reference for radiative fluxes (Loeb et al., 2009). For latent heat fluxes, we use 10 in situ, satellite-based and blended climatological products, which were analyzed in detail by (Găinuşă-Bogdan et al., 2015). References for the individual data sets are given later on in the caption of Figure 7. The da Silva et al. (1994) climatology is used for relative humidity. Cloudiness is evaluated using the CALIOP Lidar observations aboard the satellite Calipso from the Aqua-train (Chepfer et al., 2010). For comparison with models, the COSP satellite simulator is used online in the model (Bodas-Salcedo et al., 2011; Chepfer et al., 2008).

Results are shown for the reference configurations of LMDZ: versions 5A and 5B developed for CMIP5 and version 6A developed for CMIP6. These reference configurations result from long phases (typically 2 years) of careful corrections and adjustments. The 6A version includes several other changes than those of the boundary layer scheme discussed above, that are discussed by Hourdin et al. (2020). The 5A and 5B versions use an horizontal grid based on 96 equally spaced longitudes and 95 latitudes from pole to pole, and 39 vertical layers with a typical vertical resolution $\delta z = 0.3z$ in the first 3 km. The atmospheric grid of the 6A version is refined both horizontally with 144 grid points in longitude and 142 grid points in latitude, and vertically, with 79 layers and $\delta z = 0.11z$ in the first 3 km.

In order to explore the sensitivity of cloud and surface flux representation to the D parameter, a series of simulations is carried out with the 6A version by varying exclusively the value of D from 0 (without special treatment for stratocumulus) to 0.2. The names of the simulations are DXX where $XX = 100 \times D$. The standard version 6A corresponds to D07, that is, $D = 0.07$. An additional simulation named SWITCH is performed, using a threshold for the maximum temperature inversion above which the thermal plume model is intentionally switched off, as was done in 5B. In this case, the threshold value is tuned to 0.16 K/Pa to get the best agreement with the standard 6A configuration in terms of global top-of-atmosphere (TOA) radiative fluxes. The main characteristics of the various configurations are summarized in Table 1.

For coupled atmosphere-ocean simulations, we use the corresponding versions of the IPSL coupled model, IPSL-CM5A, 5B, and 6A. The results are evaluated in terms of annual mean SST using the "amip" data set.

Table 1
Characteristics of the Various LMDZ Configurations Considered in the Present Study (See Main Text for More Information)

Model version	Grid	K-diffusion	Mass flux	Treatment of stratocumulus
5A	95 × 95-L39	(Laval et al., 1981), $\Gamma = 1$ K/km	NO	K_z conditioned by a threshold on top inversion
5B	*idem*	(Yamada, 1983), $\Gamma = 0$	(Rio et al., 2010)	conditional deactivation of thermals
6A	144 × 142-L79	*idem*	*idem*	$d = f(B')$, with $D = 0.07$
DXX	*idem*	*idem*	*idem*	$d = f(B')$ with $D = 0.XX$
SWITCH	*idem*	*idem*	*idem*	conditional deactivation of thermals

Configurations 5A and 5B share the same version of the configuration ORCA2 of the oceanic model Nemo (Madec et al., 2017) with a 2° average horizontal resolution, while IPSL-CM6A uses the ORCA1 configuration, an improved version with a 1° average horizontal resolution. Two different protocols are considered for coupled simulations. For the 5A, 5B, and 6A reference configurations, we analyze directly the results of historical transient experiments run for CMIP exercises. Results are averaged over 20 consecutive years centered on year 1990 with corresponding conditions for greenhouse gases, ozone, aerosols, and land use coming from the forcing of historical simulations performed with IPSL-CM6 for CMIP6. In order to conserve computer resources and given that much of the SST bias patterns appear at short time scales (see, e.g., Vannière et al., 2014), we use 10-member ensembles of 6-month-long coupled experiments initialized in February for the sensitivity experiments to the value of the D parameter with the IPSL-CM6A configuration. The initialization of the ocean follows the protocol of the Preface project described by Voldoire et al. (2019). The 10 members differ from each other by the initial atmospheric state, which corresponds to successive 1 January taken from a reference simulation with the stand-alone atmospheric model. The first of the 10 members was integrated over 3 years.

Note that the three standard simulations were obtained at the end of a long-tuning process. It is not possible to retune in a similar way the sensitivity experiments presented here. However, a very simple retuning was done for a subset of simulations (D03T, D05T, and D10T) of these experiments, in order to restore the same global TOA net radiative fluxes, to avoid drift of the global mean temperature when coupled to ocean. This retuning was done by adjusting the coefficient that relates the aerosol concentration to the cloud droplet size (aerosols first indirect effect).

3. Results

3.1. Humidity and Clouds in Stand-Alone Atmospheric Simulations

We first analyze the near surface humidity (or relative humidity at 2 m) in the various model versions in stand-alone atmospheric simulations (Figure 1). Note that the 2 m relative humidity is a diagnostic derived from the model state variables in the first atmospheric layer and surface fluxes using similarity formulas applicable to the surface layer. It should be less dependent on the model vertical resolution than the mean relative humidity in the first atmospheric layer. In the Standard Physics version 5A, in which the boundary layer is parameterized with local turbulent diffusion only, the near-surface air humidity is globally overestimated by about 5%. With a typical observed value of relative humidity of 80%, the distance to saturation that controls surface evaporation is typically reduced by one fourth. This is due to the fact that a diffusive scheme requires a strong humidity excess in lower layers to obtain a strong upward transport. This problem could probably have been reduced by including a countergradient term on specific humidity, as is done for potential temperature.

When activating nonlocal transport by the thermal plume model in version 5B, the near-surface air dries, in better agreement with observation. However, since the mass flux scheme is switched off in regions of strong inversion, the moist bias persists over the ETO. Compared to version 5A, the root-mean-square error (RMSE, shown at the top of each panel) is reduced, from 5.6% to 4.6%, but the correlation with observation (CORR) is reduced as well, from 27% to 13%. Despite a much finer vertical resolution and other changes in the model, the SWITCH sensitivity experiments with the 6A version, which uses the same deactivation of thermal plumes in regions of strong inversion, leads to similar results (RMSE of 4.9% and correlation of 22%).

The activation of the thermal plume model everywhere in the following versions reduces the bias over the ETO as well. However, when the buoyancy of the thermal plume for detrainment is computed with respect

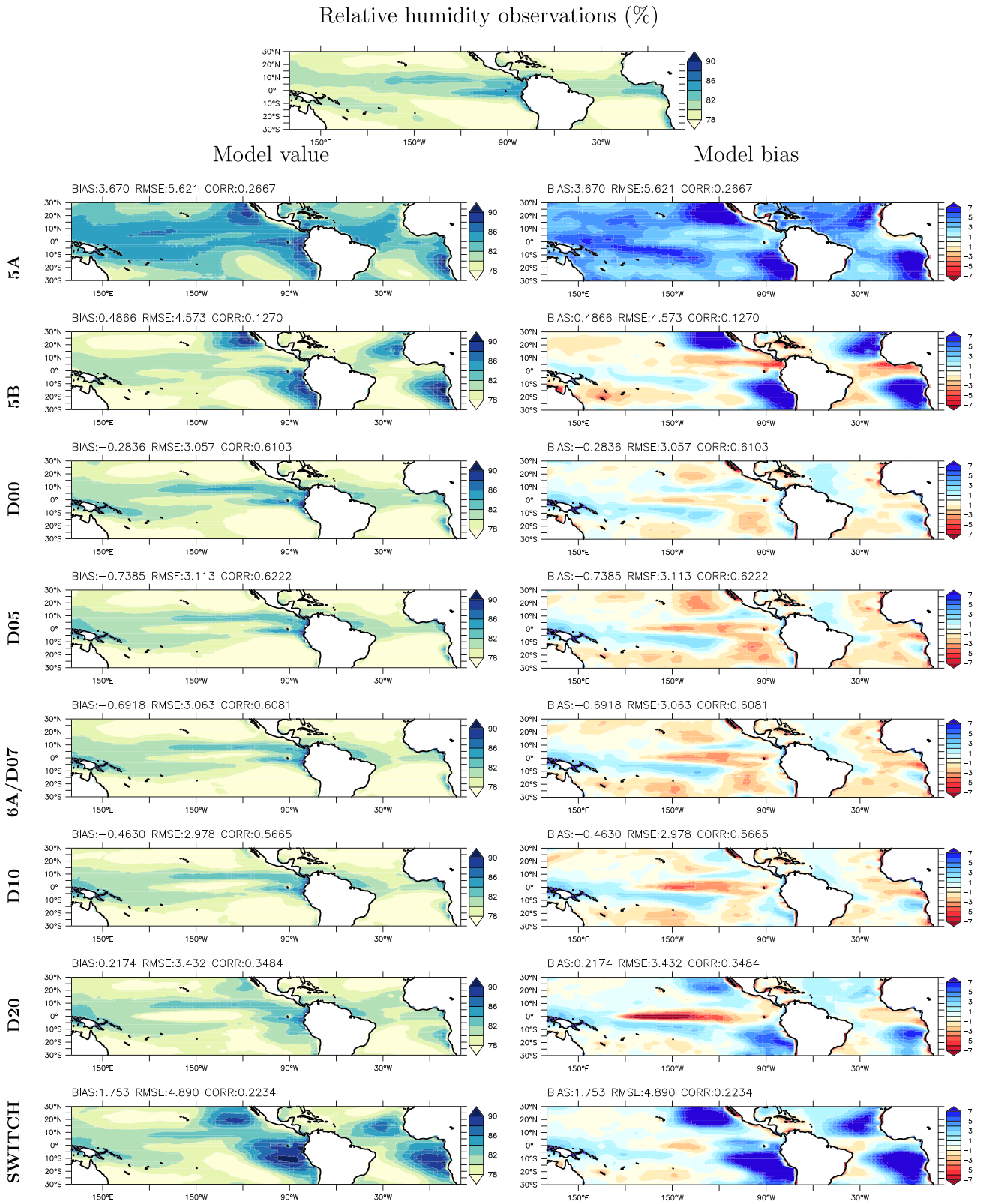

Figure 1. Annual mean near-surface relative humidity (%) in da Silva et al. (1994) observations (top panel) and in forced by SST stand-alone atmospheric simulations. The left column corresponds to model results and the right column to model biases (difference with observations). The mean bias, root-mean-square error (RMSE) and correlation (CORR) with observations are shown at the top of each panel (same for the left and right columns).

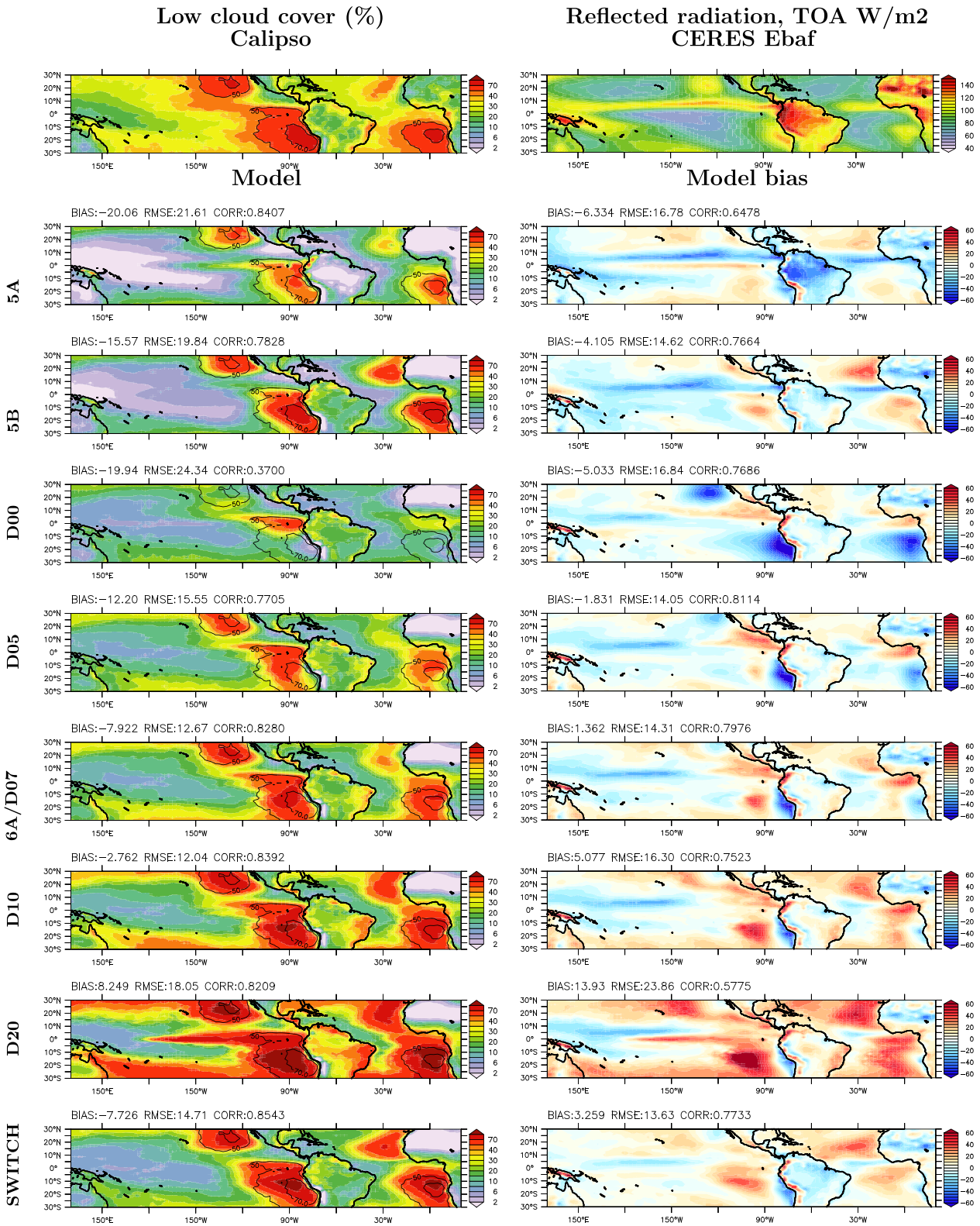

Figure 2. Low-level cloud cover (%, left) and reflected radiation at the top of the atmosphere. For clouds, we compare the calipso LIDAR observation (W m^{-2}, top) with model outputs reconstructed using the COSP calipso simulator. The contours 50% and 70% of the observations are superimposed on the color shades. For radiative fluxes, we show the model bias with respect to the ERES-EBAF climatology. The mean bias, root-mean-square error (RMSE) and correlation (CORR) with observations are shown at the top of each panel.

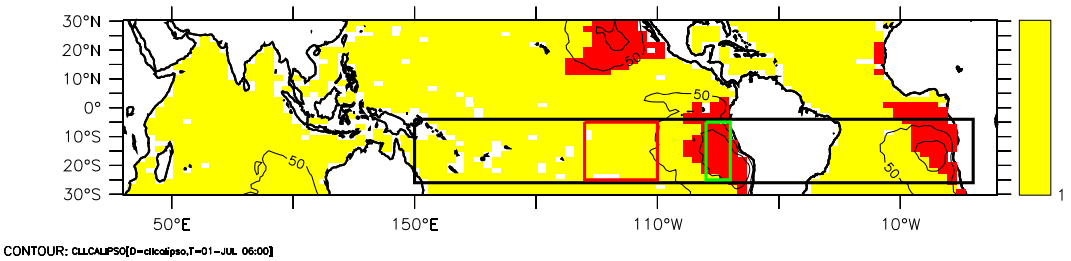

Figure 3. Masks used for analysis: transect (black rectangle), vertical profiles in the central (red), and East (green rectangle) Pacific ocean. The ETOA index is defined as the average over the red-filled area (ETO) minus the average value over the tropical oceans defined as the sum of the red and yellow areas. The 50% and 70% levels of the observed low cloud cover are shown as black contours.

to air coming from a progressively higher distance above (i.e., from simulation D00 to D20), the atmosphere progressively moistens near the surface. Simulation D20, with largest value of parameter D, shows biases that share similarities with the 5B and SWITCH simulations. The model fits the observation best for D between 0 and 0.07, showing RMSE of about 3% and correlations of about 60%. The progressive drying of the equatorial Pacific when parameter D increases probably involves interactions between the thermal plume model and the deep convection scheme and deserves further investigation.

The sensitivity to the D coefficient of the cloud cover and short-wave radiative fluxes at top of the atmosphere is shown in Figure 2. For versions 5A, 5B, and SWITCH that use a threshold on the inversion strength to switch parameterizations, the contrast between cumulus and stratocumulus regions is qualitatively captured, even if the cloud cover is systematically underestimated in the 5A version. As a result, the biases are limited between typically −30 and 30 W m^{-2}. If the thermal plume model is activated everywhere (D00), stratocumulus clouds almost disappear, leading to a bias of SW surface fluxes over the ETO that exceeds 60 W m^{-2} (associated with a similar negative bias of the reflected SW radiation). The coverage of stratocumulus clouds progressively increases with increasing values of D and becomes too strong over the ETO for $D > 0.15$, which results in a too strong reflection of solar radiation.

Again, the sensitivity experiments show modifications in other regions, For instance, over the equatorial Pacific, which deserves further investigation and may involve coupling with other processes as deep convection and large scale dynamics. The value $D = 0.07$ finally retained in LMDZ configuration 6A is a compromise, with a RMSE of 13% and a correlation of 83% for the low cloud cover and a RMSE of 14% and a correlation of 80% for radiation. Increasing D further improves cloud cover but decreases at the same time the skill for radiation. The opposite happens if decreasing the value of D. Finally, compared to simulations 5B and SWITCH where the thermal plume model is switched off when the inversion is strong, the 6A (or D07) shows similar performances in terms of cloud cover associated with a much better representation of near-surface relative humidity.

3.2. Focus on Longitudinal Contrasts

In order to further interpret the link between parameterization choices and model results, we analyze in more detail the control experiment D07 and the most extreme sensitivity experiments—D00, D20, and SWITCH—focusing on a longitudinal cross section in the southern Pacific and Atlantic oceans, averaging variables between 25°S and 5°S (black rectangle in Figure 3). Figure 4 shows the cross section of cloud cover and surface variables while Figure 5 displays vertical profiles averaged over two boxes within the Pacific part of the cross sections: a Central South Pacific box (red rectangle) and East South Pacific box (green rectangle in Figure 3).

Consistent with Figure 2, simulation D00 underestimates stratocumuls clouds (red curve, top panels in Figure 4). The D07, D20, and SWITCH simulations show a reasonable contrast between trade winds cumulus (on the West and Central oceanic basins) and stratocumulus over ETO. The vertical profiles (top panels in Figure 5) confirm this result. However, the maximum cloud coverage in SWITCH is located slightly too low compared with Calipso observations with an overestimated maximum, while the maximum cloud coverage in D07 is located too high and is a bit underestimated compared to the observations.

Consistently for D20, and moreover for SWITCH, the stratocumulus clouds are obtained at the expense of an overestimated surface relative humidity as already discussed above. The agreement with observations

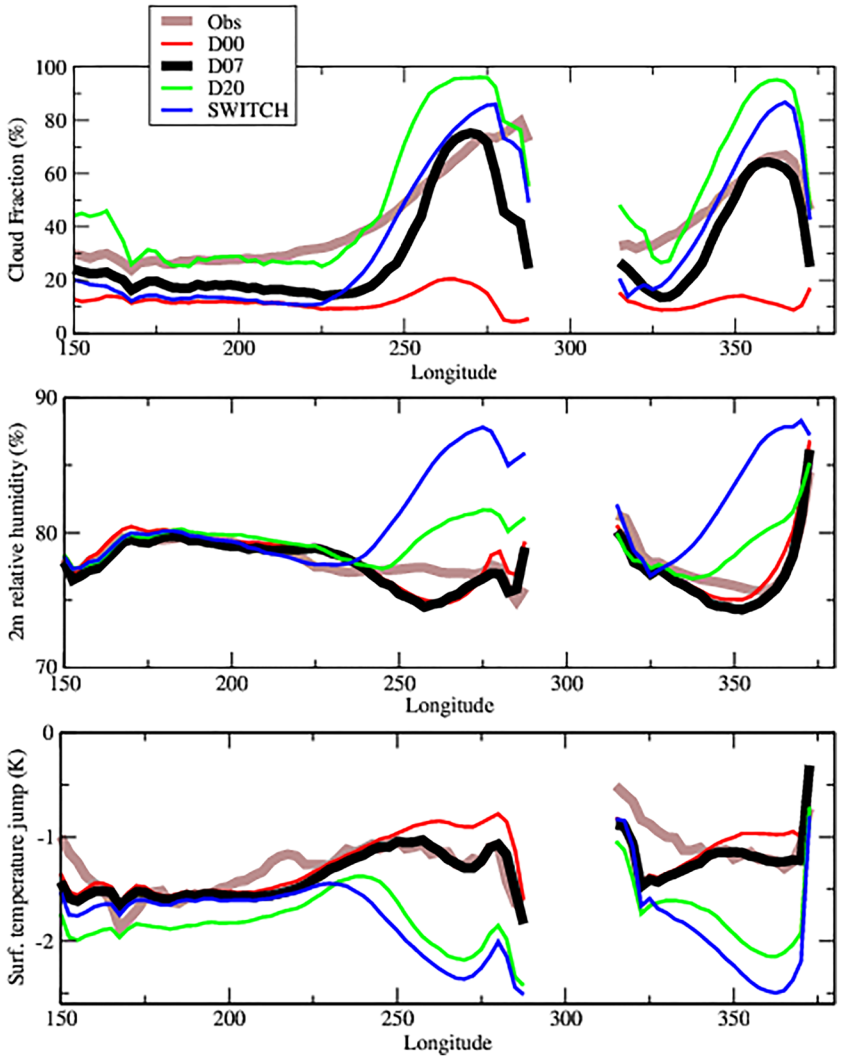

Figure 4. Cross-section average between 25°S and 15°S of the low cloud cover (top panel, %, compared with Calipso observations), 2 m relative humidity (%, middle panel, da Silva et al., 1994), and temperature contrast between the surface and near-surface air (K, bottom panel). Note that the different sets of observations used for the different graphs are completely independent of each other.

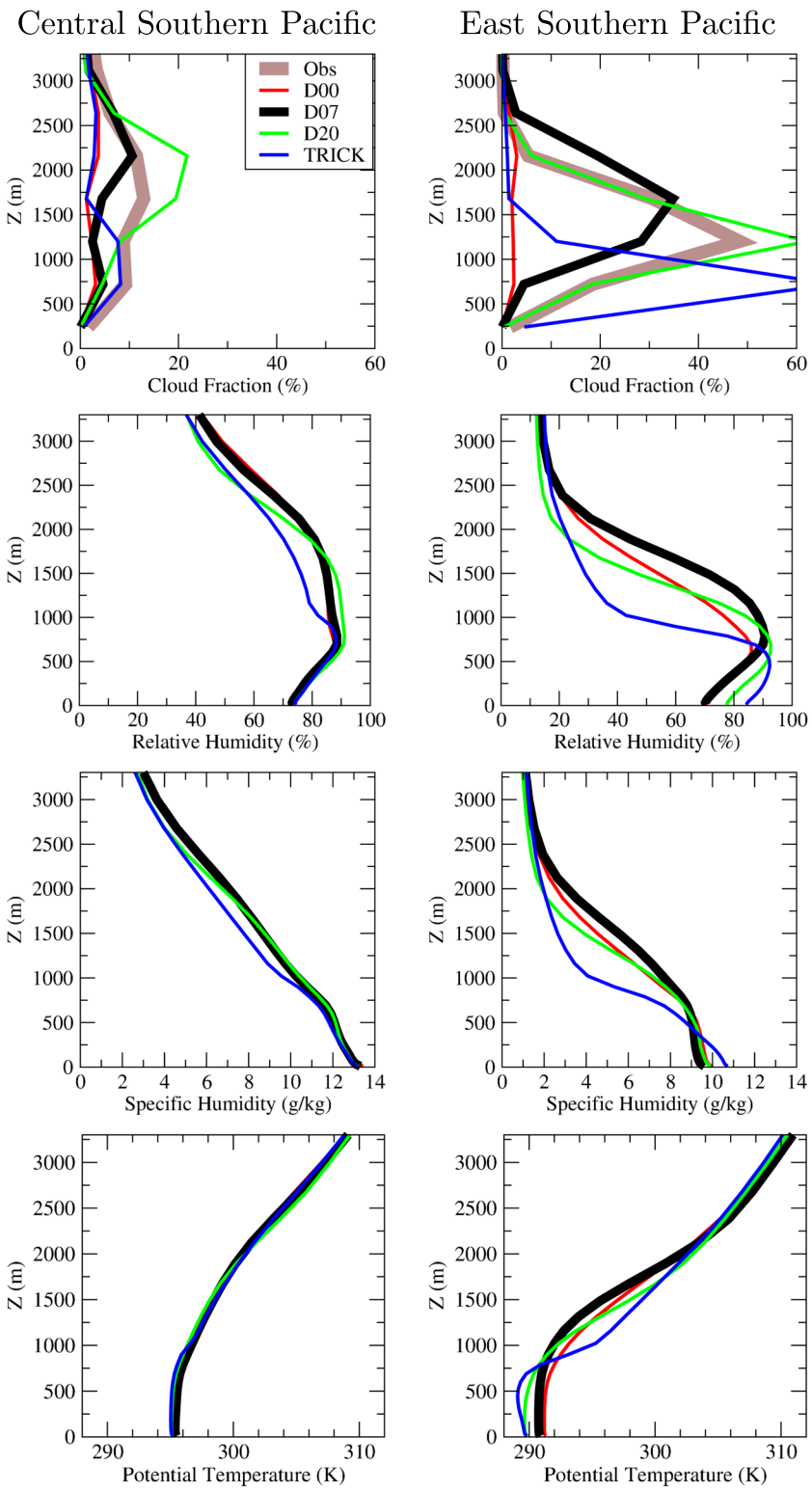

Figure 5. Vertical profiles of cloudiness (top row, %, with Calipso observations), relative humidity (second row, %), specific humidity (g/kg, third row), and potential temperature (K, bottom row) over Central Southern Pacific (left column) and East Southern Pacific (right column) in simulations D00, D07, D20, and SWITCH.

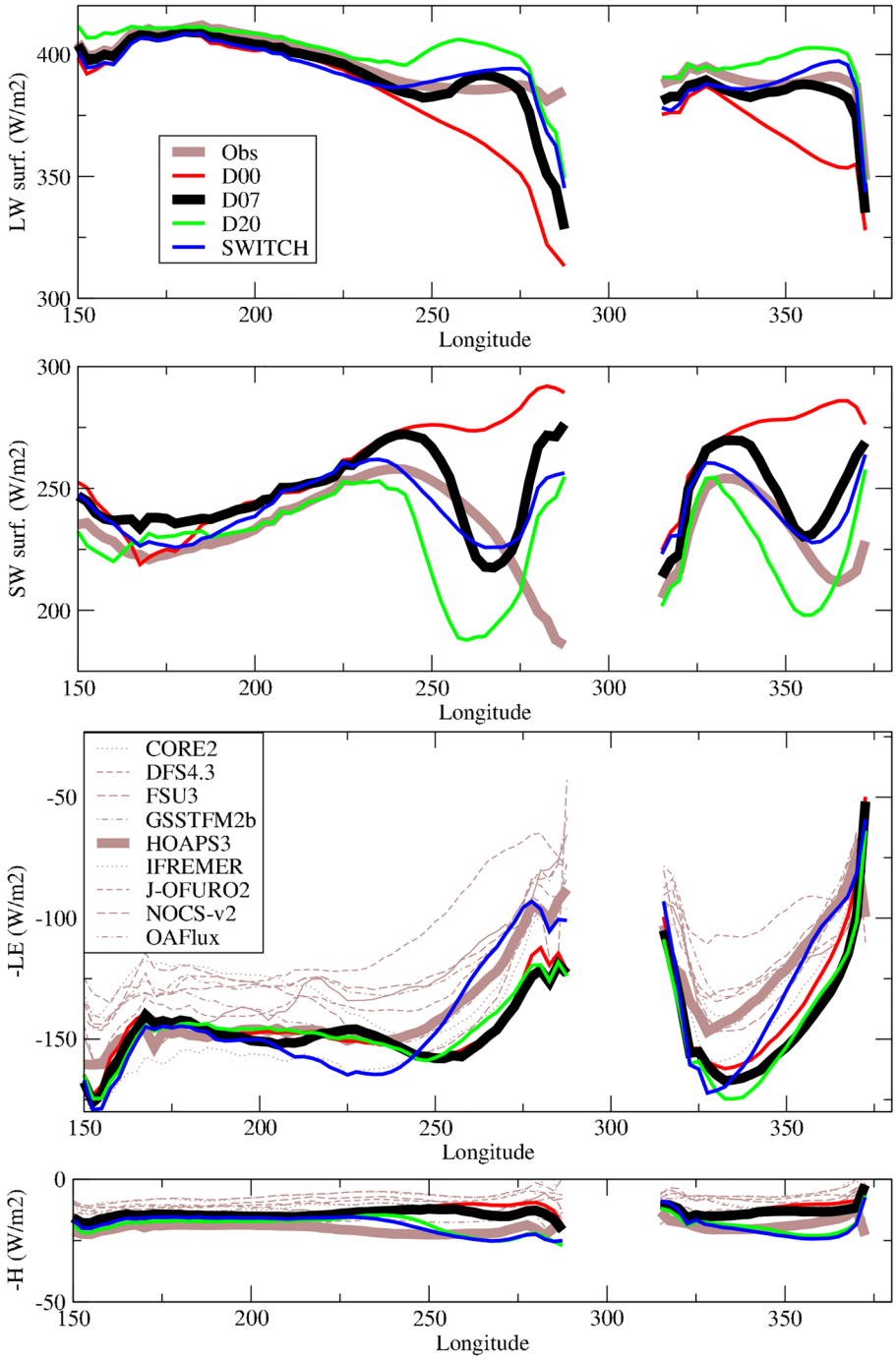

Figure 6. Cross-section average between 25°S and 5°S of the long-wave (LW) and short-wave (SW) downward surface radiation, latent and sensible heat fluxes (−LE and −H) in stand-alone atmospheric simulations (W m^{-2}) for simulations D00, D07, D20, and SWITCH.

is better for D00 and D07 in that respect. For cloudy convective boundary layers, the near-surface relative humidity is directly related to the depth of the subcloud layer, since the relative humidity is close to 100% in the cloud layer, and decreases below as a consequence of the temperature increase of 10 K/km in the dry convective subcloud layer, given that the specific humidity is almost constant in the mixed layer. The D20 and SWITCH simulations show larger relative humidity at the surface over the ETO, as a consequence of a thinner convective boundary layer. For the SWITCH simulation, the effect is reinforced by the fact that the specific humidity is not well mixed (third line in Figure 5) due to the diffusive nature of the boundary layer

transport when the thermal plume is deactivated (as already explained for simulation 5A in the previous section). Note that the link between near-surface humidity and boundary layer depth does not apply for the D00 simulations, which is far from saturation at boundary layer top.

The boundary layer in the D20 and SWITCH simulations is also colder than in D00 and D07, probably because the cloud top infrared radiative cooling is larger (the clouds are both more covering and lower—hence warmer—than in D07) and because this cooling is applied to a thinner layer. As a consequence, the change in relative humidity at the surface between D07 and D20 and the change in temperature partly compensate, resulting in a weaker change in specific humidity. To finish note that for the SWITCH simulation, the potential temperature profile is unstable (decreasing with height between the surface and 500 m), a classical result of diffusive simulations which can transport heat upward (as a convective boundary layer should do in response to upward surface sensible heat flux) only if the atmosphere is unstable.

3.3. Impact on Surface Fluxes

Clouds directly affect the surface downward radiation (see Figure 6), both in the long-wave (LW; top panel) and in the short-wave (SW; second panel). While the effect of parameterization choices is rather weak on the west of the basins, it modifies the surface energy budget over the ETO. Over ETO, D07 and SWITCH show similar results. The lack of stratocumulus clouds is visible in both the LW and SW radiation for simulation D00. In simulation D20, the LW radiation at surface is increased compared to D00, consistent with the lower altitude of cloud base and larger cloud cover.

The SWITCH simulation shows strongly reduced evaporative cooling over the ETOs (third panel). This reduced evaporative cooling is mainly a consequence of the increased humidity. However, the effect of relative humidity on evaporative cooling is not clearly visible for the simulation D20. In fact, the specific humidity at the surface is almost unchanged in the DXX simulations, probably due to the compensation between relative humidity and temperature mentioned above. In terms of agreement with observations, it is difficult to conclude which value of D is the best in view of the large dispersion among the available data sets (thin dashed curves).

Finally, the sensible heat flux is strongly reinforced over the ETO in the D20 and SWITCH simulations as a consequence of the larger temperature contrast at the surface.

Figure 7 synthetizes the results of the experiments in terms of surface fluxes by using the ETO-Anomaly index mentioned in section 1 and first introduced by Hourdin et al. (2015). The ETOA of a quantity is defined as the difference between the averaged quantity in the red area and the average over the tropical oceans 30°S to 30°N (Figure 3). The red region corresponds to the warm bias area in CMIP5 simulations. It is defined using a threshold of 0.8 K for the SST bias on the ensemble average of the CMIP5 simulations, after removal of the global mean (Hourdin et al., 2015). Note that the 70% contour of the cloud cover is almost entirely included in the ETO area. All fluxes are counted positive downward. The gray squares correspond to the total flux and the black squares to the sum of the latent heat (−LE) and Cloud Radiative Effect (CRE) in both SW and LW radiations. CRE-LE ETOA in atmospheric simulations forced by SST was shown to correlate with SST ETOA in coupled simulations (Hourdin et al., 2015). Except for the 5B version, the gray and black squares are quite close to each others suggesting that the evaporative cooling and effect of clouds dominate the intersimulation dispersion of the total surface flux. Looking at the sensitivity experiments with the 6A version, it is clear that increasing D increases the shadowing effect of clouds (blue stars in Figure 7) but that this cooling effect is in part compensated by an increased downward longwave radiation reaching the surface (green crosses in Figure 7). Using a diffusive formulation everywhere (as was the case in version 5A) or deactivating the thermals in the regions of subsidence (versions 5B and SWITCH) produces the largest ETOA for LW CRE and −LE (red circles). For values of parameter D between 0.07 and 0.1, the evaporative cooling and total fluxes reach a minimum value.

Those values correspond to a bias of -10 to $-15\ \mathrm{W\ m^{-2}}$ when compared with the observed ETOA. However, the bias is probably in the range of observational error bars as suggested by comparing the ETOA of the reference HOAPS3 data set (Andersson et al., 2010), which is usually used for evaluation of the LMDZ model, with eight alternate data sets documented by Găinuşă-Bogdan et al. (2015) (red circles on the right-hand side of the first panel of Figure 7). The NOC2 data set ($-12.6\ \mathrm{W\ m^{-2}}$, Berry & Kent, 2009) that shows the lowest −LE ETOA is much closer to the minimum of simulated ETOAs.

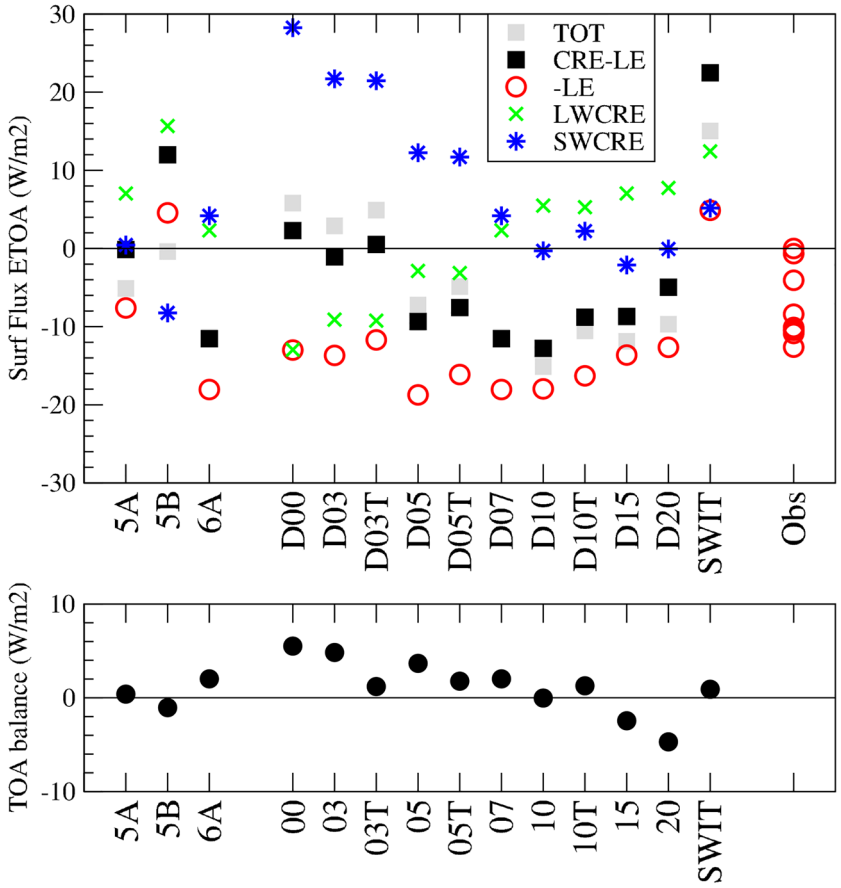

Figure 7. Bias in total surface flux ETOA (top panel) and global TOA radiative balance (bottom panel) in stand-alone atmospheric simulations. The total surface flux (TOT), the SW and LW cloud radiative forcing at surface (SWCRE and LWCRE), the latent heat (−LE), and the sum of total cloud radiative forcing (CRE=CRESW+CRELW) and latent heat (CRE-LE), as well as the TOA balance are counted positively downward. For surface fluxes, the biases are computed with respect to the reference climatological observational data set described in section 2.3. For the latent heat, the difference in ETOA with the reference HOAPS3 data set (Andersson et al., 2010) is shown in the right-hand side of the upper panel for height observational data sets analyzed by Găinuşă-Bogdan et al. (2015). Sorted with decreasing value of −LE ETOA: IFREMER (−0.7 W m^{-2}, Bentamy et al., 2003), DFS4 (−4.1 W m^{-2}, Brodeau et al., 2010), CORE2 (−8.4 W m^{-2}, Large & Yeager, 2009), FSU (−10.1 W m^{-2}, Hughes et al., 2012), GSSTF (−10.4 W m^{-2}, Shie et al., 2010), OAFlux (−10.5 W m^{-2}, Yu et al., 2011), OFURO (−10.9 W m^{-2}, Tomita et al., 2010), and NOC2 (−12.6 W m^{-2}, Berry & Kent, 2009). Note that all values are negative meaning that HOAPS3 is the data set showing the least negative value of −LE ETOA.

3.4. Impact of a Retuning of the Global Energetic Balance

With the 6A configuration, a global TOA balance of about 2 W m^{-2} in stand-alone atmospheric simulations is required in order to get a global mean surface temperature that matches observation in present-day conditions in coupled ocean simulations (in theory, this imbalance should be of 0.5 to 1 W m^{-2} corresponding to the oceanic heat uptake in the present-day warming climate). Without any retuning, a change of the TOA balance by 1 W m^{-2} would result in a change in the global mean SST of about 1 K per W m^{-2}. The modification of this balance in the various sensitivity experiments is typical of several W m^{-2} (lower panel of Figure 7), which would result in unacceptable mean biases in SSTs in coupled simulations. A common practice among modeling groups to solve this issue is to retune cloud parameters, the most uncertain parameters that affect the energetic balance the most (Hourdin et al., 2017). This was done here for simulations D03T, D05T, and D10T as explained in section 2.3.

For the three retuned simulations, as well as for the SWITCH simulation, the global radiative balance matches that of the reference configuration 6A (or D07) to less than 1 W m^{-2} (Figure 7, lower panel). In the mean time, the surface fluxes ETOA (Figure 7, upper panel) in the retuned DXXT versions do not depart from the corresponding DXX version by more than 3 W m^{-2}, which is small compared to the sensitivity to

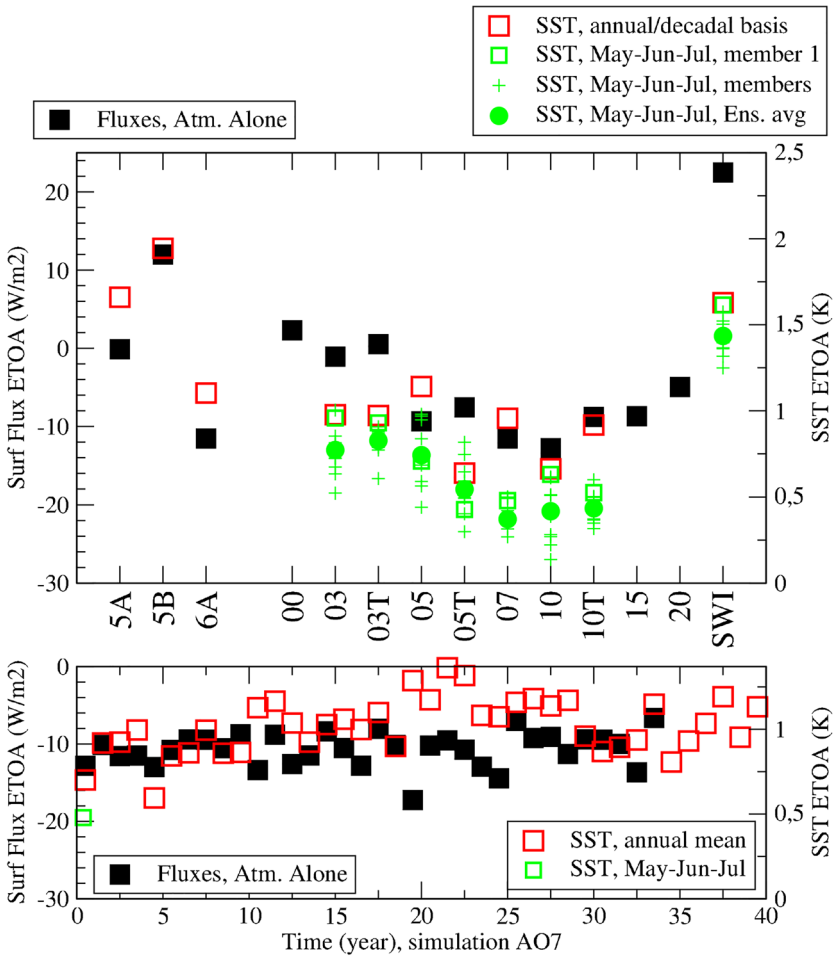

Figure 8. SST ETOA in coupled to ocean simulations compared with the surface fluxes ETOA in the forced by SST simulations (black squares, same as in Figure 7). For the coupled simulations with IPSL-CM5A, 5B and 6A versions, we show averaged SSTs between 1980 and 2000 (red squares). For the sensitivity experiments, the red squares correspond to the average over the first year of the first member of the 10-member ensemble. The full green circles correspond to the average over the months of May, June, and July and over the 10 members of the ensemble. Individual member values are shown by a plus sign, except for member #1 which corresponds to the full green square. For the reference simulation with version 6A, the year-to-year variation of the flux ETOA in stand-alone atmospheric simulations is shown on the lower panel, as well as the average over May-June-July of the first year, and SST ETOA in coupled simulation.

the value of D. Thus, a slight retuning of the global energetic balance, such as the one applied when deriving a configuration of the coupled model, does not affect the conclusions concerning the ETOA of surface fluxes and SSTs (next section).

3.5. Effect of the Scheme Modification on SST in Coupled Simulations

Coupled simulations were run for a subset of the configurations documented above in stand-alone atmospheric mode: the three reference CMIP configurations of IPSL-CM (5A, 5B, and 6A) as well as a subset of the sensitivity experiments with the 6A configuration.

For simulations 5A, 5B, and 6A, we analyze directly the corresponding CMIP historical simulations for the present-day period. Averages are done over 20 consecutive years centered on 1990. The bias in SST ETOA index is shown in Figure 8 (right-hand side axis) together with the associated biases in fluxes ETOA in stand-alone atmospheric simulations (left axis, black squares, already in Figure 7). The 1.7 K SST ETOA bias of the IPSL-CM5A version, which was typical of CMIP5 models was in fact even larger in the 5B version (1.9 K), due to smaller evaporative cooling over the ETO, where the thermal plume model was switched off, and

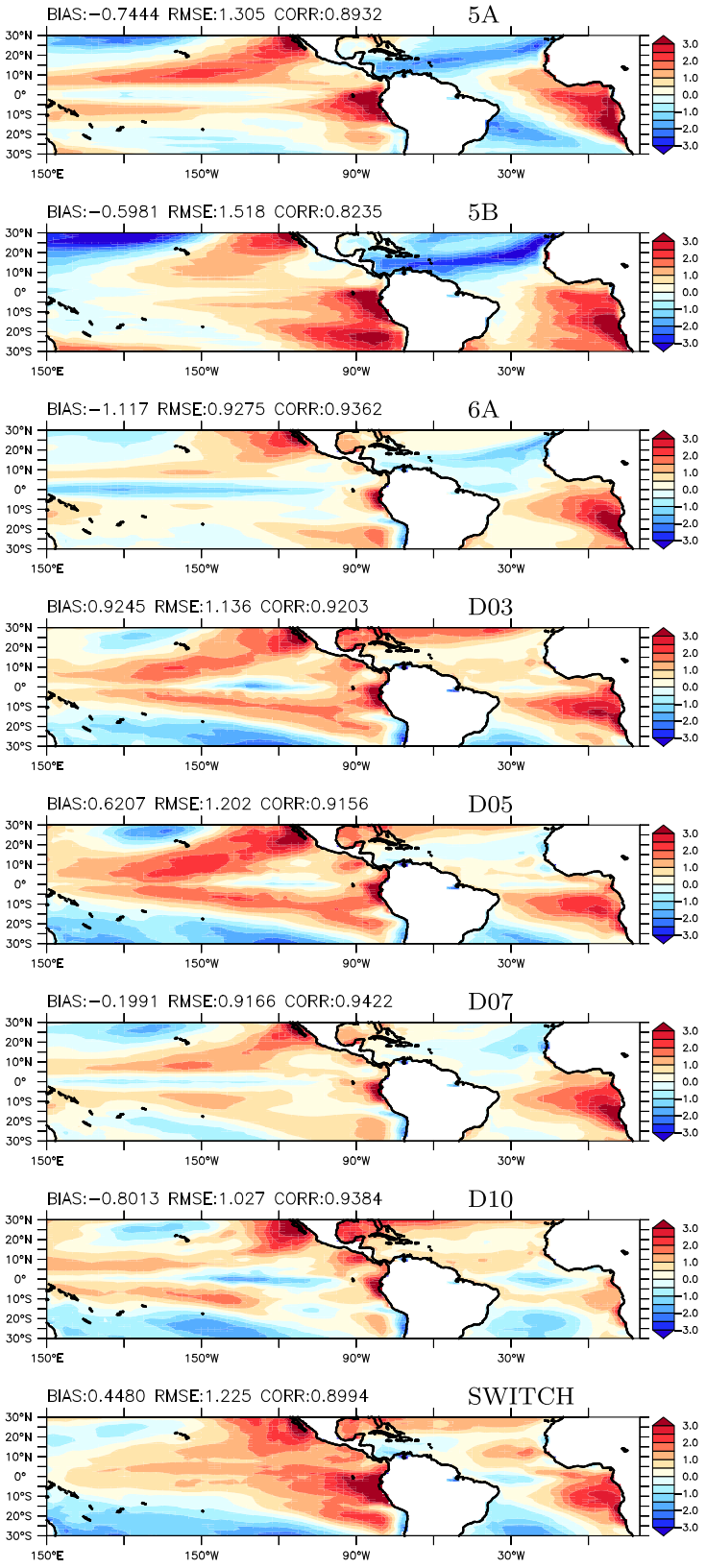

Figure 9. Annual mean SST bias, mean value removed. For the reference CMIP configuration (top three panels), the average over year 1980 to 2000 is shown for the first member of "historical" simulations. For the other panels, the third year of the simulation is shown.

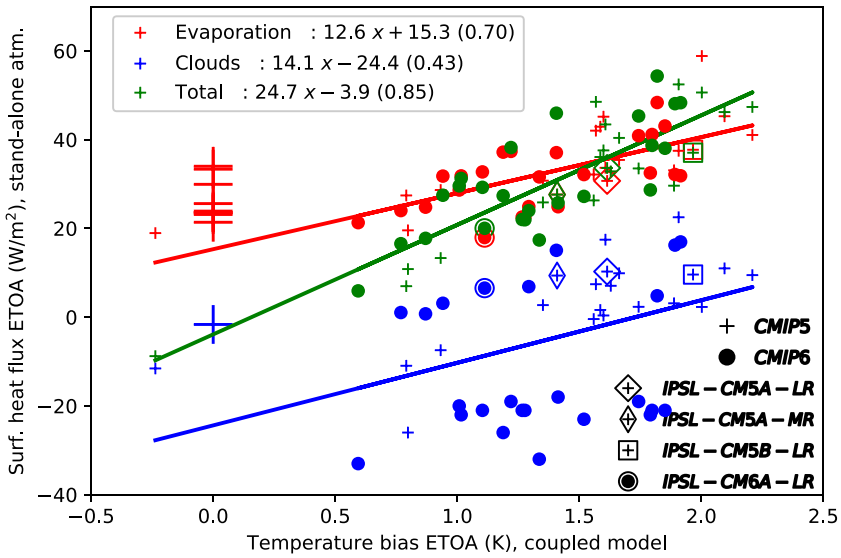

Figure 10. Scatter plot of atmospheric fluxes ETOA in stand-alone amip simulations as a function of the SST ETOA bias in coupled simulations for the same model. Fluxes are shown for cloud radiative effect (blue), for latent heat (red), and for the total surface heat flux (i. e., total radiation plus latent and sensible heat fluxes). Equations of the regression lines (full lines) and correlation coefficients are given in the legend. The red and blue large plus signs at 0 on the x axis show the individual observational estimates (same as in Figure 7). The regression lines are computed using all the available models for which historical coupled simulations and amip stand-alone simulations were available on IPSL extraction of the CMIP database. CMIP5 models are shown as small plus signs and the CMIP6 models as full circles. The IPSL configurations are encircled by empty markers, as shown in the lower right part of the graph. The list of CMIP models sorted as a function of the SST ETOA, together with the values of surface fluxes are given in Tables A1 and A2.

to an increased downward LW radiation because of clouds closer to the surface. As was seen for the fluxes, the bias is significantly reduced in the new 6A version, although it does not disappear.

Regarding the sensitivity experiments with the 6A version, we use the 10-member ensemble of short-term initialized simulations described in section 2.3.

The green plus signs in Figure 8 correspond to the average over May-June-July for these 10 members. For the first member, which was run over 3 years, the May-June-July average is shown as a green open square and the average over the third year as a red square. The ensemble average of the 10 members is shown as a full green circle for May-June-July. There is a rather large intermember dispersion, of typically 0.5 K for a given configuration between the largest and smallest ETOA. The year-to-year variability is also of the order of 0.5 K and that of the flux ETOA is of typically 10 W m^{-2} as seen in the lower panel of Figure 8 (there is also a slight trend in the SST ETOA). Despite this internal variability, there is a quite consistent response of the SST ETOA to the surface fluxes. In particular, the D07 and D10 simulations show minimum values of the SST ETOA bias and the SWITCH simulation the largest consistently with fluxes in simulations forced by SST, for the same reason as the 5B version.

As expected, the bias is larger after 3 years than between months 4 and 6 after initialization. Values of the D coefficient around 0.1 lead to small ETOA SST biases, while a strong bias is obtained for the SWITCH simulation. When looking in more detail, the D07 simulation gives the smallest SST bias for the ensemble average after 6 months for the initialized simulations. Note, however, that the interannual variability of this ETOA index for SST is large, as illustrated in the lower panel of Figure 8, consistently with the large intermember dispersion.

Figure 9 shows the maps of SST biases, corresponding to the red squares in Figure 8. Compared with versions 5A and 5B, the SST bias is almost reduced everywhere in version 6A. Note, however, that the reduction of the warm bias is not as strong over the south Atlantic and north Pacific than over the south Pacific basin, a point which deserves additional analysis. The D10 simulation is the only one showing a strong reduction of this particular bias, but to the expense of somewhat larger biases in other areas compared to simulation D07. This result may be partly due to the internal variability since only a 3 year simulation is available for D10.

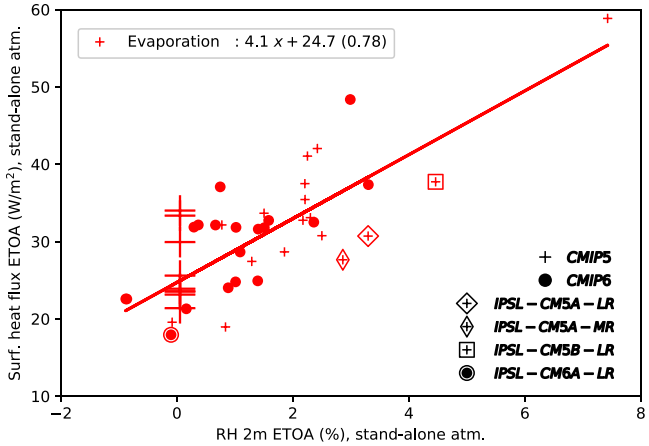

Figure 11. Scatter plot of atmospheric latent fluxes ETOA in stand-alone amip simulations as a function of the near-surface relative humidity ETOA bias in the same simulations. Same conventions as in Figure 10 concerning the markers code.

Note that it is not the aim of this study to fix once and for all the best value of parameter *D* but rather to restrain its range of possible values to enter it in the process of model tuning for the future versions of the IPSL coupled model.

Figure 10 presents the comparison of the successive versions of the CMIP configurations of the IPSL model with other CMIP models. Figure 10 is an update with CMIP6 simulations of Figure 1e of Hourdin et al. (2015). The scatter plot shows a good correlation (of 0.70 for latent heat, 0.43 for radiation, and 0.85 for the total) between fluxes ETOA in stand-alone atmospheric simulations (on the *y* axis) and SST ETOA bias (on the *x* axis). Except for the CSIRO CMIP5 model, all the CMIP5 and CMIP6 models show SST ETOA bias of more than 0.6 K. A variation of the total flux ETOA of about 28 W m^{-2} is required to gain 1 K in the SST ETOA bias. The fact that the blue and red regression lines intercept 0 on the *x* axis for values of the flux ETOA smaller than observations (shown as large plus signs) may either suggest a bias in observations or be an indication that about 20 W m^{-2} in the forcing of the SST ETOA bias comes from a lack of cooling in the oceanic model (either due to a deficiency of the model itself or to errors in the atmospheric drags).

The degradation of the IPSL model from the 5A to 5B version in terms of ETOA is clearly visible, as well as an improvement when increasing the horizontal resolution between the 5A-LR and 5A-MR versions. With a SST ETOA bias of 1.1 K, the ISPL-CM6A-LR version is better than the previous versions. This improvement is obtained together with an improvement of the physics. In particular, the contrasted behavior over the ETO and the rest of the tropics is obtained without switching parameterizations on or off. Note that part of this improvement may also come from the increased oceanic resolution, an issue which would require to test the two oceanic resolutions with the exactly same atmospheric model. Some models are doing better than IPSL-CM6A-MR in terms of SST ETOA. Further investigation is needed to determine if it is due to a better physical behavior, to a finer grid, or to compensating errors. Once again, the uncertainty in observed surface fluxes probably prevents to be more conclusive on the model that performs best in that respect.

4. Discussion

By extending previous analysis to the new CMIP6 multi-model ensemble, the present study first confirms conclusions reached by Hourdin et al. (2015) and Găinuşă-Bogdan et al. (2018): (1) the warm SST bias over the ETO is due in a large part to erroneous representation of the contrast between surface fluxes over the trade winds and stratocumulus regions; (2) not only the shadowing effect of clouds is important but also their greenhouse effect; (3) the evaporative cooling plays a role as large as radiation. Here, we enlighten the link between errors in surface fluxes and key aspects of atmospheric parameterizations of turbulence, boundary layer convection, and cloud processes. The parameterization must be able to maintain a strong contrast between trade winds cumulus regions with cloud cover of about 10% and the ETO where stratocumulus clouds cover very large areas during all seasons. The simulated cloud base height of stratocumulus clouds is also important as a cloud base too close to the surface induces an overestimation of the cloud greenhouse effect. Moreover, an adequate representation of the latent heat flux requires a good representation of boundary layer convection. In LMDZ, the activation of the thermal plume model, replacing the counter-gradient term, dries the surface by exporting more efficiently the air evaporated at the surface to the free dry troposphere or by importing more efficiently dry tropospheric air to the surface.

The positive correlation between overestimated relative humidity and underestimated evaporation in terms of ETOA in the CMIP5 and CMIP6 multimodel ensemble (Figure 11) points to a control of surface evaporative cooling by atmospheric boundary layer processes. The dispersion of the representation of the ETOA of humidity and latent heat flux by itself points to a misrepresentation of those processes in at least part of the models. If the fix suggested here for the particular physical parameterizations of LMDZ may not be applicable to other models, this work could motivate other modeling centers to pay closer attention to the representation of nonlocal boundary layer mixing and entrainment at boundary layer top and its impact on regional fluxes and SSTs. This issue is clearly challenging for parameterizations which should represent

both the control of near-surface humidity by convection over a layer, which is 500 m to 1 km deep and the maintenance of a very thin cloud at its top with a quasi-discontinuity at its upper boundary. It may partly explains the long-standing difficulty of coupled model to represent SSTs well in this area.

The approach retained for the 6A version of LMDZ, a mass flux representation of the vertical transport in the convective boundary layer by the thermal plume model, was shown, both at process scale and in global simulation, to provide several improvements compared to a diffuse approach, including (1) the countergradient transport of heat in convective goundary layer (Hourdin et al., 2002); (2) the diurnal cycle of near-surface wind over deserts (Hourdin et al., 2015); (3) the diurnal cycle of clouds and near-surface humidity over continents (Diallo et al., 2017; Rio & Hourdin, 2008); (4) the vertical transport of atmospheric tracers (Locatelli et al., 2015), and more recently, after the introduction of the modification of lateral detrainment, (5) the representation of top entrainment of water vapor and transition from stratocumulus to cumulus clouds (Hourdin et al., 2019). Note that the introduction of the D parameter, although representing a significant step forward both by avoiding switching parameterizations on or off depending on a critical temperature gradient and by providing a reasonable representation of the boundary layer top entrainment, would probably deserve to be replaced by a more physical representation of the entrainment of dry air from the free troposphere into the boundary layer, as, for example, by an explicit representation of downdrafts within stratocumulus clouds (Brient et al., 2019). Due to the lack of relevant observations or too much uncertainty in the products available, it is difficult to say with certainty wether the contrast in the vertical transport of moisture between ETO and the rest of the tropics is improved in version 6A compared to previous versions. Compared with ERA-Interim reanalysis (not shown), LMDZ6A shows a moister lower troposphere suggesting a stronger entrainment at boundary layer top. However, given the short time scale of the turbulent and convective transport that controls vertical humidity profiles in the first km above surface, given the scarcity of observation and the error in surface turbulent fluxes in ERA-Interim when compared to buoy observations (see Table 1 in Zuidema et al., 2016), ERA-Interim cannot be considered as a reference for this aspect.

Table A1
List of CMIP5 Model Configurations Analyzed in Figures 10 and 11

Model	SST bias (K)	Latent ($W\ m^{-2}$)	Sensible ($W\ m^{-2}$)	SW CRE ($W\ m^{-2}$)	LW CRE ($W\ m^{-2}$)	TOTAL ($W\ m^{-2}$)	RH (%)
CSIRO-Mk3-6-0	−0.237	18.951	−11.552	−19.9405	−8.464	−8.7779	0.838
ACCESS1-3	0.791	27.432	−10.954	−23.7355	−6.331	6.9877	1.2909
CESM1-CAM5	0.798	19.567	−26.567	−9.8869	−9.227	10.8416	−0.0823
ACCESS1-0	0.932	28.658	−7.439	−18.3255	−8.76	13.3278	1.8546
CCSM4	1.351	30.756	2.723	−3.2804	−12.4	25.8627	2.4999
IPSL-CM5A-MR	1.409	27.644	9.377	−10.372	−11.031	27.618	2.8612
MRI-CGCM3	1.559	32.145	−0.421	−9.9997	−11.589	26.3142	0.7755
GISS-E2-R	1.568	42.029	7.43	6.3416	−7.702	48.525	2.422
MPI-ESM-LR	1.586	42.9144	1.6216	−5.4892	−9.256	36.0927	—
MPI-ESM-MR	1.599	45.1748	0.3832	−6.7032	−8.054	37.5871	—
MIROC5	1.607	33.663	17.458	13.7836	−10.599	43.4335	1.5023
IPSL-CM5A-LR	1.614	30.71	10.277	−9.84	−11.653	33.6146	3.297
GFDL-CM3	1.629	32.752	7.068	1.8198	−13.872	32.8508	2.1738
inmcm4	1.663	35.433	9.897	−4.5889	−6.504	40.3939	2.2088
bcc-csm1-1	1.743	40.2456	2.3304	−6.3576	−9.702	33.5157	—
NorESM1-M	1.888	33.0904	3.1286	−3.8399	−11.922	29.6265	2.3016
CNRM-CM5	1.907	37.49	22.516	29.1914	−15.385	52.4527	2.2068
IPSL-CM5B-LR	1.966	37.726	9.603	−15.8564	−11.725	37.0587	4.4609
FGOALS-s2	2.002	58.8534	2.2996	−1.9289	−5.433	50.6291	7.4263
FGOALS-g2	2.095	45.2822	9.7178	2.5962	−8.149	46.2164	—
MIROC-ESM	2.209	41.0533	9.4807	6.6737	−10.354	47.3665	2.2509

Note. The models are ranked according to the SST ETOA bias (column 2). The other columns show the ETOA index for surface fluxes and radiative effects and relative humidity.

Table A2
List of CMIP6 model configurations analysed in Figures 10 and 11

Model	SST bias (K)	Latent ($W\ m^{-2}$)	Sensible ($W\ m^{-2}$)	SW CRE ($W\ m^{-2}$)	LW CRE ($W\ m^{-2}$)	TOTAL ($W\ m^{-2}$)	RH (%)
SAM0-UNICON	0.594	21.303	−33.303	−18.7125	−9.098	5.9274	0.1621
CESM2-WACCM	0.769	24.022	1.046	−4.6002	−8.511	16.5567	0.8848
CESM2	0.87	24.786	0.754	−5.1304	−8.033	17.7651	1.0114
UKESM1-0-LL	0.941	31.791	3.135	−3.0505	−9.417	27.4774	1.5073
GFDL-CM4	1.008	28.652	−19.652	−2.5426	−8.877	29.5928	1.0898
FGOALS-f3-L	1.016	31.836	−22.836	−5.6885	−6.676	31.2597	1.0192
HadGEM3-GC31-LL	1.103	32.735	−21.735	−2.181	−9.944	29.2567	1.5822
IPSL-CM6A-LR	1.112	17.943	6.548	−8.7883	−8.221	20.0292	−0.1011
BCC-CSM2-MR	1.188	37.187	−26.187	−8.9544	−9.09	27.3727	—
FGOALS-g3	1.22	37.3587	−18.3587	1.1778	−11.436	38.2239	3.3011
EC-Earth3-Veg	1.266	22.57	−21.57	−5.1328	−11.001	22.0121	−0.8632
EC-Earth3	1.277	22.58	−21.58	−5.2395	−10.984	21.9493	−0.8815
MRI-ESM2-0	1.293	24.916	6.881	−6.8676	−8.07	24.0077	1.3933
NorESM2-LM	1.336	31.605	−31.605	−15.5362	−5.7	17.3899	1.4026
MIROC6	1.406	37.071	15.062	12.2796	−11.269	45.9682	0.7469
E3SM-1-0	1.413	24.888	−18.888	0.4084	−8.038	25.7659	—
CanESM5	1.519	32.135	−24.135	−6.2803	−11.17	27.2228	0.665
NESM3	1.743	40.8978	−18.8978	−2.5367	−14	45.3341	—
NorCPM1	1.791	32.506	−21.506	−4.3447	−11.807	28.6751	2.3605
BCC-ESM1	1.798	41.1638	−20.1638	−1.7488	−11.39	38.7126	—
GISS-E2-1-G	1.819	48.3686	4.8204	−0.031	−7.455	54.3334	2.9894
CAMS-CSM1-0	1.851	43.0763	−21.0763	−4.2869	−8.672	38.0618	—
CNRM-ESM2-1	1.892	32.15	16.242	11.6959	−10.375	48.1099	0.3692
CNRM-CM6-1	1.915	31.863	16.957	12.4688	−10.437	48.3284	0.293

Note. The models are ranked according to the SST ETOA bias (column 2). The other columns show the ETOA index for surface fluxes and radiative effects and relative humidity.

The fact that the reduction of SST biases was obtained together with a better agreement with the available climatology for relative humidity as well as the better representation of boundary layer processes and clouds documented in both 1-D and 3-D by Hourdin et al. (2019) are however strong arguments suggesting that the representation of boundary layer mixing processes was indeed improved.

Note that the 6A configuration of the model was obtained after a long and iterative tuning process, the ETOA contrast in fluxes (enforced by SST atmospheric simulations) or SST (in coupled simulations) being one target of the tuning process. Estimating the part of the other parameters and processes that control the strength of the ETOA fluxes and SSTs is however not feasible with the current hand-made way of exploring the space of parameters for tuning. More automatic and systematic tools should be used to achieve this, as metamodels or emulators proposed in the frame of the Uncertainty Quantification field. Such approaches are currently under development in the LMDZ team.

Even with a flux ETOA bias negative by 10 $W\ m^{-2}$, a warm ETOA bias persists in SSTs however. As was already noted by Hourdin et al. (2015) the fact that the correct representation of SSTs in terms of ETOA index requires a negative value of the surface fluxes, ETOA bias suggests that the inconsistency could come from the observations. The discrepancy would be reduced if using other observational data sets than the HOAPS3, which is the less favorable for the comparison with LMDZ latent heat flux. It is also possible that the tuning of the current version of LMDZ6A overfits the SSTs by pushing the negative value of the surface flux ETOA too far to compensate for a lack of oceanic cooling, due to errors in the oceanic model, or in the wind stress.

These results advocate for significant efforts to constrain surface variables over tropical oceans from observations.

Appendix A: CMIP5 and CMIP6 Values of the ETOA Index

Tables A1 and A2 show lists of CMIP5 and CMIP6 model configurations analyzed in Figures 10 and 11.

Acknowledgments
This work was supported by the DEPHY2 project, funded by the French national program LEFE/INSU. The 3-D simulations were granted access to the HPC resources of IDRIS and TGCC under the allocation gencmip6 attributed by GENCI (Grand Equipement National de Calcul Intensif). This study benefited from the ESPRI computing and data centre (https://mesocentre.ipsl.fr) which is supported by CNRS, Sorbonne Université, Ecole Polytechnique, and CNES as well as through national and international grants. The work benefited from a support of the FP7 "Preface" project (Grant 603521). The last version of the LMDZ source code can be downloaded freely from the LMDZ web site. The version used for the specific simulations runs for this paper is the "svn" release 3404 from 16 October 2018, which can be downloaded and installed on a Linux computer by running the "install_lmdz.sh" script available at this site (http://www.lmd.jussieu.fr/lmdz/pub:./install_lmdz.sh. -v20181016.trunk). A large part of the output used are archived on the CMIP5 and CMIP6 archives, distributed through the Earth System Grid Federation (ESGF) and freely accessible through the ESGF data portals after registration. Details about ESGF are presented on the CMIP Panel website (http://www.wcrp-climate.org/index.php/wgcm-cmip/about-cmip). The multi-model analysis benefited as well from the effort of the Convergence ANR poject (grant# ANR-13-MONU-0008). Preprocessed files will be made available on a DOI if the paper is accepted for publication, together with the scripts used to generate the figures.

References

Andersson, A., Fennig, K., Klepp, C., Bakan, S., Graßl, H., & Schulz, J. (2010). The Hamburg Ocean atmosphere parameters and fluxes from satellite data—HOAPS-3. *Earth System Science Data*, *2*, 215–234. https://doi.org/10.5194/essd-2-215-2010

Arakawa, R. A., & Schubert, W. H. (1974). Interaction of a cumulus cloud ensemble with the large scale environment. Part I. *Journal of the Atmospheric Sciences*, *31*, 674–701.

Bentamy, A., Katsaros, K. B., Mestas-Nuñez, A. M., Drennan, W. M., Forde, E. B., & Roquet, H. (2003). Satellite estimates of wind speed and latent heat flux over the global oceans. *Journal of Climate*, *16*, 637–656. https://doi.org/10.1175/1520-0442(2003)016<0637:SEOWSA>2.0.CO;2

Berry, D. I., & Kent, E. C. (2009). A new air-sea interaction gridded dataset from ICOADS with uncertainty estimates. *Bulletin of the American Meteorological Society*, *90*, 645. https://doi.org/10.1175/2008BAMS2639.1

Bodas-Salcedo, A., Webb, M. J., Bony, S., Chepfer, H., Dufresne, J. L., Klein, S. A., & John, V. O. (2011). COSP: Satellite simulation software for model assessment. *Bulletin of the American Meteorological Society*, *92*(8), 1023–1043. https://doi.org/10.1175/2011BAMS2856.1

Boucher, O., Servonnat, J., Albright, A.-L., Aumont, O., Balkanski, Y., Bastrikov, V., et al. (2020). Presentation and evaluation of the IPSL-CM6A-LR climate. *Journal of Advances in Earth Modeling Systems*, *12*. https://doi.org/10.1029/2019MS002010

Bretherton, C. S., & Park, S. (2009). A new moist turbulence parameterization in the community atmosphere model. *Journal Climate*, *22*(12), 3422–3448. https://doi.org/10.1175/2008JCLI2556.1

Brient, F., Couvreux, F., Villefranque, N., Rio, C., & Honnert, R. (2019). Object-oriented Identification of coherent structures in large eddy simulations: Importance of downdrafts in stratocumulus. *Geophysical Research Letters*, *46*, 2854–2864. https://doi.org/10.1029/2018GL081499

Brodeau, L., Barnier, B., Treguier, A. M., Penduff, T., & Gulev, S. (2010). An ERA40-based atmospheric forcing for global ocean circulation models. *Ocean Modelling*, *31*, 88–104. https://doi.org/10.1016/j.ocemod.2009.10.005

Chatfield, R. B., & Brost, R. A. (1987). A two-stream model of the vertical transport of trace species in the convective boundary layer. *Journal of Geophysical Research*, *92*, 13,263–13,276.

Chepfer, H., Bony, S., Winker, D., Cesana, G., Dufresne, J. L., Minnis, P., & Zeng, S. (2010). The GCM Oriented Calipso Cloud Product (CALIPSO-GOCCP). *Journal of Geophysical Research*, *115*, D00H16. https://doi.org/10.1029/2009JD012251

Chepfer, H., Bony, S., Winker, D., Chiriaco, M., Dufresne, J. L., & Sèze, G. (2008). Use of CALIPSO lidar observations to evaluate the cloudiness simulated by a climate model. *Geophysical Research Letters*, *351*, L15704. https://doi.org/10.1029/2008GL034207

Cheruy, F., Campoy, A., Dupont, J. C., Ducharne, A., Hourdin, F., Haeffelin, M., & Idelkadi, A. (2013). Combined influence of atmospheric physics and soil hydrology on the simulated meteorology at the SIRTA atmospheric observatory. *Climate Dynamics*, *40*, 2251–2269. https://doi.org/10.1007/s00382-012-1469-y

D'Orgeval, T., Polcher, J., & de Rosnay, P. (2008). Sensitivity of the West African hydrological cycle in ORCHIDEE to infiltration processes. *Hydrology and Earth System Sciences*, *12*(6), 1387–1401.

da Silva, A., Young, A. C., & Levitus, C. (1994). Atlas of surface marine data 1994, volume 1: Algorithms and procedures. 6, U.S. Department of Commerce, NOAA, NESDIS.

de Szoeke, S. P., Fairall, C. W., Wolfe, D. E., Bariteau, L., & Zuidema, P. (2010). Surface flux observations on the southeastern Tropical Pacific Ocean and attribution of SST errors in coupled ocean-atmosphere models. *Journal Climate*, *23*, 4152–4174. https://doi.org/10.1175/2010JCLI3411.1

Deardorff, J. W. (1966). The counter-gradient heat-flux in the lower atmosphere and in the laboratory. *Journal of the Atmospheric Sciences*, *23*, 503–506.

Diallo, F. B., Hourdin, F., Rio, C., Traore, A. K., Mellul, L., Guichard, F., & Kergoat, L. (2017). The surface energy budget computed at the grid-scale of a climate model challenged by station data in West Africa. *Journal of Advances in Modeling Earth Systems*, *9*(7), 2710–2738. https://doi.org/10.1002/2017MS001081

Emanuel, K. A. (1991). A scheme for representing cumulus convection in large-scale models. *Journal of the Atmospheric Sciences*, *48*, 2313–2335.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S., Collins, W., & Rummukainen, M. (2013). Evaluation of climate models. In T. Stocker (Ed.), *Climate change 2013: The physical scientific basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 741–866). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. https://doi.org/10.1017/CBO9781107415324.020

Găinuşă-Bogdan, A., Hourdin, F., Traore, A. K., & Braconnot, P. (2018). Omens of coupled model biases in the CMIP5 AMIP simulations. *Climate Dynamics*, *51*, 2927–2941. https://doi.org/10.1007/s00382-017-4057-3

Găinuşă-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. *Journal of Geophysical Research: Atmospheres*, *120*, 4483–4505. https://doi.org/10.1002/2014JD022985

Grandpeix, J., & Lafore, J. (2010). A density current parameterization coupled with Emanuel's convection scheme. Part I: The models. *Journal of Atmospheric Sciences*, *67*, 881–897. https://doi.org/10.1175/2009JAS3044.1

Grandpeix, J., Lafore, J., & Cheruy, F. (2010). A density current parameterization coupled with Emanuel's convection scheme. Part II: 1D simulations. *Journal of Atmospheric Sciences*, *67*, 898–922. https://doi.org/10.1175/2009JAS3045.1

Holtslag, A. A. M., & Boville, B. A. (1993). Local versus non-local boundary-layer diffusion in a global climate model. *Journal Climate*, *6*, 1825–1842.

Hourdin, F., Couvreux, F., & Menut, L. (2002). Parameterisation of the dry convective boundary layer based on a mass flux representation of thermals. *Journal of the Atmospheric Sciences*, *59*, 1105–1123.

Hourdin, F., Foujols, M. A., Codron, F., Guemas, V., Dufresne, J. L., Bony, S., & Bopp, L. (2013). Impact of the LMDZ atmospheric grid configuration on the climate and sensitivity of the IPSL-CM5A coupled model. *Climate Dynamics*, *40*, 2167–2192. https://doi.org/10.1007/s00382-012-1411-3

Hourdin, F., Găinuşă-Bogdan, A., Braconnot, P., Dufresne, J. L., Traore, A. K., & Rio, C. (2015). Air moisture control on ocean surface temperature, hidden key to the warm bias enigma. *Geophysical Research Letters*, *42*, 10,885–10,893. https://doi.org/10.1002/2015GL066764

Hourdin, F., Grandpeix, J. Y., Rio, C., Bony, S., Jam, A., Cheruy, F., & Roehrig, R. (2013). LMDZ5B: the atmospheric component of the IPSL climate model with revisited parameterizations for clouds and convection. *Climate Dynamics*, *40*, 2193–2222. https://doi.org/10.1007/s00382-012-1343-y

Hourdin, F., Gueye, M., Diallo, B., Dufresne, J. L., Escribano, J., Menut, L., & Guichard, F. (2015). Parameterization of convective transport in the boundary layer and its impact on the representation of the diurnal cycle of wind and dust emissions. *Atmospheric Chemistry and Physics*, *15*, 6775–6788. https://doi.org/10.5194/acp-15-6775-2015

Hourdin, F., Jam, A., Rio, C., Couvreux, F., Sandu, I., Lefebvre, M. P., & Idelkadi, A. (2019). Unified parameterization of convective boundary layer transport and clouds with the thermal plume model. *Journal of Advances in Modeling Earth Systems*, *11*, 2910–2933. https://doi.org/10.1029/2019MS001666

Hourdin, F., Mauritsen, T., Gettelman, A., Golaz, J. C., Balaji, V., Duan, Q., & Williamson, D. (2017). The art and science of climate model tuning. *Bulletin of the American Meteorological Society*, *98*, 589–602. https://doi.org/10.1175/BAMS-D-15-00135.1

Hourdin, F., Musat, I., Bony, S., Braconnot, P., Codron, F., Dufresne, J. L., & Lott, F. (2006). The LMDZ4 general circulation model: climate performance and sensitivity to parametrized physics with emphasis on tropical convection. *Climate Dynamics*, *27*, 787–813. https://doi.org/10.1007/s00382-006-0158-0

Hourdin, F., Rio, C., Grandpeix, J.-Y., Madeleine, J.-B., Cheruy, F., Rochetin, N., et al. (2020). LMDZ6A: The atmospheric component of the IPSL climate model with improved and better tuned physics. *Journal of Advances in Modeling Earth Systems*, *12*, e2019MS001892. https://doi.org/10.1029/2019MS001892

Hughes, P. J., Bourassa, M. A., Rolph, J. J., & Smith, S. R. (2012). Averaging-related biases in monthly latent heat fluxes. *Journal of Atmospheric and Oceanic Technology*, *29*, 974–986. https://doi.org/10.1175/JTECH-D-11-00184.1

Jam, A. (2012). Paramétrisation de la couche limite: Du modèle des thermiques au modèle statistique de nuage (PhD thesis), advized by Hourdin, Frédéric, Sciences de l'environnement, Paris 6 2012, 2012PA066672. https://www.theses.fr/2012PA066672

Jam, A., Hourdin, F., Rio, C., & Couvreux, F. (2013). Resolved versus parametrized boundary-layer plumes. Part III: Derivation of a statistical scheme for cumulus clouds. *Boundary-Layer Meteorology*, *147*, 421–441. https://doi.org/10.1007/s10546-012-9789-3

Large, W. G., & Yeager, S. G. (2009). The global climatology of an interannually varying air-sea flux data set. *Climate Dynamics*, *33*, 341–364. https://doi.org/10.1007/s00382-008-0441-3

Laval, K., Sadourny, R., & Serafini, Y. (1981). Land surface processes in a simplified general circulation model. *Geophysical & Astrophysical Fluid Dynamics*, *17*, 129–150.

Locatelli, R., Bousquet, P., Hourdin, F., Saunois, M., Cozic, A., Couvreux, F., & Williams, A. G. (2015). Atmospheric transport and chemistry of trace gases in LMDz5B: evaluation and implications for inverse modelling. *Geoscientific Model Development*, *8*, 129–150. https://doi.org/10.5194/gmd-8-129-2015

Lock, A. P. (2001). The numerical representation of entrainment in parameterizations of boundary layer turbulent mixing. *Monthly Weather Review*, *129*(5), 1148. https://doi.org/10.1175/1520-0493(2001)129<1148:TNROEI>2.0.CO;2

Loeb, N. G., Wielicki, B. A., Doelling, D. R., Smith, G. L., Keyes, D. F., Kato, S., & Wong, T. (2009). Toward optimal Ccosure of the Earth's top-of-atmosphere radiation budget. *Journal Climate*, *22*(3), 748–766. https://doi.org/10.1175/2008JCLI2637.1

Ma, C. C., Mechoso, C. R., Robertson, A. W., & Arakawa, A. (1996). Peruvian stratus clouds and the tropical Pacific circulation: A coupled ocean-atmosphere GCM study. *Journal of Climate*, *9*, 1635–1645. https://doi.org/10.1175/1520-0442(1996)009<1635:PSCATT>2.0.CO;2

Madec, G., Bourdallé-Badie, R., Bouttier, P., Bricaud, C., Bruciaferri, D., Calvert, D., & Vancoppenolle, M. (2017). NEMO ocean engine (Version v3.6). Notes du Ple de modélisation de l'Institut Pierre-simon Laplace (IPSL), https://doi.org/10.5281/zenodo.1472492

Mellor, G. L., & Yamada, T. (1974). A hierarchy of turbulence closure models for planetary boundary layers. *Journal of the Atmospheric Sciences*, *31*, 1791–1806.

Richter, I. (2015). Climate model biases in the eastern tropical oceans: Causes, impacts and ways forward. *WIREs Climate Change*, *6*, 345–358. https://doi.org/10.1002/wcc.338

Rio, C., & Hourdin, F. (2008). A thermal plume model for the convective boundary layer: Representation of cumulus clouds. *Journal of the Atmospheric Sciences*, *65*, 407–425.

Rio, C., Hourdin, F., Couvreux, F., & Jam, A. (2010). Resolved versus parametrized boundary-layer plumes. Part II: Continuous formulations of mixing rates for mass-flux schemes. *Boundary-Layer Meteorology*, *135*, 469–483. https://doi.org/10.1007/s10546-010-9478-z

Rio, C., Hourdin, F., Grandpeix, J., & Lafore, J. (2009). Shifting the diurnal cycle of parameterized deep convection over land. *Geophysical Research Letters*, *36*, L07809. https://doi.org/10.1029/2008GL036779

Rochetin, N., Grandpeix, J. Y., Rio, C., & Couvreux, F. (2014). Deep convection triggering by boundary layer thermals. Part II: Stochastic triggering parameterization for the LMDZ GCM. *Journal of the Atmospheric Sciences*, *71*, 515–538. https://doi.org/10.1175/JAS-D-12-0337.1

Shie, C., Chiu, L., Adler, R., Gao, S., Chokngamwong, R., Lin, I., & Wang, F. (2010). A recently revived and produced global air-sea surface turbulent fluxes dataset-GSSTF2b: Validations and findings. In *Proceedings of the Joint 2010 CWB Weather Analysis and Forecasting and COAA 5th International Ocean-Atmosphere Conference*.

Siebesma, A. P., Soares, P. M. M., & Teixeira, J. (2007). A combined eddy-diffusivity mass-flux approach for the convective boundary layer. *Journal of Atmospheric Sciences*, *64*, 1230–1248. https://doi.org/10.1175/JAS3888.1

Soares, P. M. M., Miranda, P. M. A., Siebesma, A. P., & Teixeira, J. (2004). An eddy-diffusivity/mass-flux parametrization for dry and shallow cumulus convection. *Quarterly Journal of the Royal Meteorological Society*, *130*, 3365–3383.

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*, 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Taylor, K. E., Williamson, D., & Zwiers, F. (2000). The sea surface temperature and sea-ice concentration boundary conditions for AMIP II simulations. PCMDI report 60, Lawrence Livermore National Laboratory, Livermore, California, Program for Climate Model Diagnosis and Intercomparison.

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. *Journal of Geophysical Research*, *115*, C03018. https://doi.org/10.1029/2009JC005545

Vannière, B., Guilyardi, E., Toniazzo, T., Madec, G., & Woolnough, S. (2014). A systematic approach to identify the sources of tropical SST errors in coupled models using the adjustment of initialised experiments. *Climate Dynamics*, *43*, 2261–2282. https://doi.org/10.1007/s00382-014-2051-6

Vignon, E., Hourdin, F., Genthon, C., Van de Wiel, B. J. H., Gallée, H., Madeleine, J. B., & Beaumet, J. (2018). Modeling the dynamics of the atmospheric boundary layer over the Antarctic Plateau with a general circulation model. *Journal of Advances in Modeling Earth Systems*, *10*, 98–125. https://doi.org/10.1002/2017MS001184

Voldoire, A., Exarchou, E., Sanchez-Gomez, E., Demissie, T., Deppenmeier, A. L., Frauen, C., & Traoré, A. K. (2019). Role of wind stress in driving SST biases in the Tropical Atlantic. *Climate Dynamics*, *53*, 3481–3504. https://doi.org/10.1007/s00382-019-04717-0

Xu, Z., Chang, P., Richter, I., Kim, W., & Tang, G. (2014). Diagnosing southeast tropical Atlantic SST and ocean circulation biases in the CMIP5 ensemble. *Climate Dynamics*, *43*, 3123–3145. https://doi.org/10.1007/s00382-014-2247-9

Xu, Z., Li, M., Patricola, C. M., & Chang, P. (2014). Oceanic origin of southeast tropical Atlantic biases. *Climate Dynamics*, *43*, 2915–2930. https://doi.org/10.1007/s00382-013-1901-y

Yamada, T. (1983). Simulations of nocturnal drainage flows by a q^2l turbulence closure model. *Journal of the Atmospheric Sciences*, *40*, 91–106.

Yu, B., Boer, G. J., Zwiers, F. W., & Merryfield, W. J. (2011). Covariability of SST and surface heat fluxes in reanalyses and CMIP3 climate models. *Climate Dynamics*, *36*, 589–605. https://doi.org/10.1007/s00382-009-0669-6

Yu, J. Y., & Mechoso, C. R. (1999). Links between annual variations of Peruvian stratocumulus clouds and of SST in the eastern equatorial Pacific. *Journal Climate*, *12*, 3305–3318. https://doi.org/10.1175/1520-0442(1999)012<3305:LBAVOP>2.0.CO;2

Zheng, Y., Shinoda, T., Lin, J. L., & Kiladis, G. N. (2011). Sea surface temperature biases under the stratus cloud deck in the Southeast Pacific Ocean in 19 IPCC AR4 coupled general circulation models. *Journal Climate*, *24*, 4139–4164. https://doi.org/10.1175/2011JCLI4172.1

Zuidema, P., Chang, P., Medeiros, B., Kirtman, B. P., Mechoso, R., Schneider, E. K., & Xu, Z. (2016). Challenges and prospects for reducing coupled climate model SST biases in the eastern tropical Atlantic and Pacific Oceans: The U.S. CLIVAR eastern tropical oceans synthesis working group. *Bulletin of the American Meteorological Society*, *97*(12), 2305–2327. https://doi.org/10.1175/BAMS-D-15-00274.1