

# Variations of Tropical Lapse Rates in Climate Models and Their Implications for Upper-Tropospheric Warming

P. KEIL,<sup>a,b</sup> H. SCHMIDT,<sup>a</sup> B. STEVENS,<sup>a</sup> AND J. BAO<sup>a</sup>

<sup>a</sup> *Max-Planck Institute for Meteorology, Hamburg, Germany*

<sup>b</sup> *International Max Planck Research School on Earth System Modelling, Hamburg, Germany*

(Manuscript received 10 March 2021, in final form 8 September 2021)

**ABSTRACT:** The vertical temperature structure in the tropics is primarily set by convection and therefore follows a moist adiabat to first order. However, tropical upper-tropospheric temperatures differ among climate models and observations, as atmospheric convection remains poorly understood. Here, we quantify the variations in tropical lapse rates in CMIP6 models and explore reasons for these variations. We find that differences in surface temperatures weighted by the regions of strongest convection cannot explain these variations and, therefore, we hypothesize that the representation of convection itself and associated small-scale processes are responsible. We reproduce these variations in perturbed physics experiments with the global atmospheric model ICON-A, in which we vary autoconversion and entrainment parameters. For smaller autoconversion values, additional freezing enthalpy from the cloud water that is not precipitated warms the upper troposphere. Smaller entrainment rates also lead to a warmer upper troposphere, as convection and thus latent heating reaches higher. Furthermore, we show that according to most radiosonde datasets all CMIP6 AMIP simulations overestimate recent upper-tropospheric warming. Additionally, all radiosonde datasets agree that climate models on average overestimate the amount of upper-tropospheric warming for a given lower-tropospheric warming. We demonstrate that increased entrainment rates reduce this overestimation, likely because of the reduction of latent heat release in the upper troposphere. Our results suggest that imperfect convection parameterizations are responsible for a considerable part of the variations in tropical lapse rates and also part of the overestimation of warming compared to the observations.

**SIGNIFICANCE STATEMENT:** A major criticism of climate model simulations has been their overestimation of warming in the tropical upper troposphere, between 8- and 13-km altitude, compared to observations. We show that climate models already disagree on the mean upper-tropospheric temperatures, even before warming. We demonstrate that the process of how much a convective cloud mixes with its surroundings, so-called entrainment, significantly influences upper-tropospheric temperatures and their rate of warming. Increasing entrainment decreases the heat released by condensation, which in turn reduces upper-tropospheric warming to resemble the observed warming. Improving the representation of this process in climate models, as well as other aspects of convection, should therefore be beneficial for the simulation of upper-tropospheric temperatures.

**KEYWORDS:** Atmosphere; Tropics; Atmospheric circulation; Clouds; Convection; Deep convection; Entrainment; Climate change; Cumulus clouds; Latent heating/cooling; Radiosonde/rawinsonde observations; Convective parameterization; General circulation models

## 1. Introduction

Air parcels undergoing deep convection change their temperature during their ascent according to the moist adiabatic lapse rate. In the tropics any horizontal temperature gradients produced by deep convection are quickly reduced by gravity waves (Bretherton and Smolarkiewicz 1989), resulting in a fairly weak temperature gradient (WTG). Thus, the lapse rate throughout the tropical troposphere is set by deep convection and follows a moist adiabat closely (Stone and Carlson 1979).

However, this is a simplified picture, and neglects some crucial details. For starters, the effect of entrainment is important, since undiluted ascent is very rare (Romps and Kuang 2010), and entrainment has been shown to influence upper-tropospheric stratification by regulating latent heating in the convecting plumes (Singh and O’Gorman 2013). Also, the

lapse rate is likely not set by the single warmest and deepest convective plume, but rather a spectrum of entraining convective plumes (Zhou and Xie 2019; Bao and Stevens 2021). Further, it is not clear to what degree the ascent follows an idealized moist pseudoadiabat, which assumes instant removal of condensate (all cloud water precipitates), or a reversible moist adiabat, which assumes no removal of condensate at all (no precipitation), or something in between (Bao and Stevens 2021). Another aspect to be considered is the fusion enthalpy, which is a source of cloud buoyancy (Romps and Kuang 2010). Finally, at some level in the upper troposphere, the radiative-convective equilibrium starts to transition to a purely radiative equilibrium (Folkins 2002). Since the WTG approximation holds reasonably well, and the mean tropical lapse rate is indeed primarily set by deep convection (Bao and Stevens 2021), all of these processes should have an influence on the mean observed lapse rate in the tropics. Due to deficient resolutions climate models usually parameterize many of these processes, and do so in a range of different ways (Plant and Yano 2016).

Corresponding author: P. Keil, paul.keil@mpimet.mpg.de

DOI: 10.1175/JCLI-D-21-0196.1

© 2021 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy ([www.ametsoc.org/PUBSReuseLicenses](http://www.ametsoc.org/PUBSReuseLicenses)).

It should be noted that the WTG assumption applies to the virtual temperature and thus drier regions of the tropical troposphere should be slightly warmer.

Uncertainties have also been reported in relation to global warming. Under greenhouse gas forcing, the tropical upper troposphere is expected to warm more than the surface and lower troposphere, since cloud base saturation vapor pressure is a strongly increasing function of temperature, and this additional vaporization enthalpy is realized by a disproportionate warming with height (Santer et al. 2005). However, the observed warming in the early twenty-first century is significantly weaker than predicted by climate models and basic theory (Santer et al. 2005; Thorne et al. 2007; Mitchell et al. 2013; Fu et al. 2011; Santer et al. 2017b,a; Suárez-Gutiérrez et al. 2017), although results depend on the exact time period (Thorne et al. 2007; Suárez-Gutiérrez et al. 2017), and observations also hold uncertainties (Sherwood et al. 2005; Thorne et al. 2007, 2011; Po-Chedley et al. 2015). It has been suggested that deficiencies in the post-2000 forcing (Santer et al. 2017b), as well as a wrong representation of SSTs and their coupling to deep convection (Flannaghan et al. 2014; Fueglistaler et al. 2015; Tuel 2019) might impact upper-tropospheric warming rates. Indeed, the bias is smaller in atmospheric models that use observed SSTs, compared to coupled atmosphere–ocean models (Mitchell et al. 2013; Po-Chedley et al. 2021). Furthermore, entrainment dampens the warming of the tropical troposphere by reducing the additional vaporization enthalpy (Singh and O’Gorman 2013) which is likely one reason why the overestimation of warming by climate models is not as drastic as expected from the theoretical adiabats (Miyawaki et al. 2020).

Increasing the conceptual understanding of what processes determine the tropical upper-tropospheric lapse rate and reducing these uncertainties in climate models could be beneficial for the representation of many other aspects of global circulation and climate. For example, the strength of the Walker circulation and its evolution under greenhouse gas forcing (Sohn et al. 2016), the atmospheric moisture flux into the Arctic (Lee et al. 2019) as well as tropical cyclone intensity (Trabing et al. 2019) have been shown to depend on the tropical upper-tropospheric stratification. Because tropical mid- to upper-tropospheric temperature affects baroclinity in the midlatitudes it also impacts midlatitude eddies and poleward heat transport (Lu and Cai 2010; Wu et al. 2011). In addition, the response of tropical anvil clouds to greenhouse gas warming likely depends on upper-tropospheric static stability [proportionally higher anvil temperature hypothesis (PHAT); Zelinka and Hartmann (2010)], and may result in a cloud feedback that impacts equilibrium climate sensitivity.

In this study we investigate the diverse representations of tropical lapse rates across climate models. In section 3 we document differences in the mean lapse rates and upper-tropospheric temperatures among phase 6 of the Coupled Model Intercomparison Project (CMIP6) models. We find that precipitation weighted sea surface temperatures (PR SSTs; Fueglistaler et al. 2015; Tuel 2019) do not explain the variation of upper-tropospheric temperatures in the mean state better than lower-tropospheric temperatures. Therefore, in section 4 we demonstrate how differences in the way climate models

parameterize precipitating deep convection itself has a large influence on upper-tropospheric temperatures. We do this by changing parameters in the convection and microphysics parameterizations in AMIP experiments with the atmospheric component of the climate model ICON-ESM (hereafter, ICON-A). Finally, in section 5 we examine recent upper-tropospheric warming in our ICON-A experiments, CMIP6 models and radiosonde observations and investigate how the warming from the lower troposphere is transferred to the upper troposphere.

## 2. Methods

### a. CMIP6

To study variations in upper-tropospheric temperatures we use the preindustrial control (piControl) and the Atmospheric Model Intercomparison Project (AMIP) experiments of CMIP6 (Eyring et al. 2016). In section 3a we analyze tropical lapse rates in the CMIP6 piControl and AMIP simulations. In sections 3b and 3c we use the AMIP simulations to compare to the ICON experiments and analyze historical warming rates, respectively. In the piControl experiments the climate is equilibrated, which is ideal to study the time mean properties of tropical lapse rates. In contrast, the AMIP experiments are forced by observed sea surface temperatures and therefore do not represent the stationary state (i.e., the mean state), but provide a somewhat realistic framework to analyze historical warming. Also, the warming trend should be less conflated with uncertainties arising from internal variability compared to a coupled simulation (e.g., a RCP scenario), in which SSTs vary substantially (Mitchell et al. 2013). We use the first ensemble member of every model (“r1i1p1f1”) and the entire available timespan, which differs from model to model in the piControl case, and the years 1979–2012 in the AMIP case, since some radiosonde products do not extend beyond 2012. Fifty-one models provide air temperature, and of those 50 provide surface temperature and precipitation in the piControl case. In the AMIP case 40 models provide air temperature, while 38 models provide air temperature, surface temperature and precipitation.

### b. Observations

Recent studies analyzing upper-tropospheric temperatures have primarily made use of satellite data, which provides spatially complete measurements in the tropics. Radiosonde data have been used less frequently, perhaps because the spatial coverage is sparse and they have been suggested to be error prone (Sherwood et al. 2005; Thorne et al. 2011). However temperatures inferred from satellites have also been shown to underestimate tropical tropospheric warming rates and have been continuously corrected (Po-Chedley et al. 2015). Radiosonde data have the advantage of higher vertical resolution, and therefore we make use of various different radiosonde products in this study to analyze tropical lapse rates. As we will show, the radiosondes products are overall in reasonable agreement among each other and with the ERA5, which strengthens our confidence in the radiosonde data.

For the analysis of the mean state in sections 3 and 4, we use the Iterative Universal Kriging version 2 (IUKv2; Sherwood and Nishant 2015) radiosonde dataset to compare models with observations. The IUKv2 addresses many of the issues identified as limitations in the past, as it considers time-changing instrument biases. Unlike the other radiosonde products, it provides absolute temperatures, which makes it suitable to assess the time mean tropical lapse rates. However, it provides no estimate of a tropical average, only data for individual stations. We use data from 69 stations in the tropics (20°N–20°S) over the 1979–2014 period. Tropical means that are shown refer to simple averaging over all stations. When subsampling the model data to the grid points of the radiosonde locations, results in section 3a are very similar.

For the analysis of upper-tropospheric warming in section 5 we also include radiosonde data from various other sources, namely, the HADAT (Thorne et al. 2005); RATPAC (Free et al. 2005); Rich-obs, Rich-tau, and Raobcore (Haimberger et al. 2012); and SUNY-Albany (Zhou et al. 2021) datasets. These datasets provide temperature anomalies as a tropical mean (20°N–20°S) or as gridded data, from which we calculate the tropical mean. Homogenization over time, as in the IUKv2 case, is also applied in different manners for all of these products. We show perturbed parametric ensemble estimates for Rich-obs and Rich-tau until the year 2012, which indicate the range of uncertainty in these products.

In addition to the radiosondes we also use the ERA5 (Hersbach et al. 2020) as an estimate for the lapse rate and upper-tropospheric warming. For air temperature in years 2000–06 we use ERA5.1 that shows more realistic stratospheric temperatures (Simmons et al. 2020).

### c. ICON-A experiments and convection parameterization

We employ the atmospheric component of the ICON Earth System Model ICON-A (Giorgetta et al. 2018) to investigate the sensitivity of the tropical lapse rate to perturbations in the convection scheme. We choose the R2B4 AMIP configuration (160-km horizontal grid spacing and 47 vertical levels) for various reasons: ICON-A produces a realistic climate in the tropics in this configuration (Crueger et al. 2018), the resolution is representative of those in the CMIP6 ensemble and it is computationally feasible to perform many experiments over the AMIP time period (1979–2014) to investigate the role of different parameters and also account for internal variability. We simulate a reference climate with seven ensemble members and experiments with perturbed autoconversion and turbulent entrainment with five ensemble members for each parameter value. The chosen parameter values are given in Table 1.

The convection parameterization in ICON-A is based on the scheme proposed by Tiedtke (1989) with some adaptations by Nordeng (1994). A comprehensive summary of the scheme as implemented in ICON-A is given by Möbis and Stevens (2012). It is based on the bulk equations that calculate the convective mass flux  $M_u$  as a function of homogeneous entrainment  $E_u$  and detrainment  $D_u$ . The subscript  $u$  denotes the updraft variables:

$$\frac{\partial M_u}{\partial z} = E_u - D_u. \quad (1)$$

TABLE 1. Parameter variations in the ICON-A perturbed physics experiments.

	Entrainment $\epsilon_u$	Autoconversion $K$
Reference	$2 \times 10^{-4}$	$2.5 \times 10^{-4}$
Low	$0.4 \times 10^{-4}$	$0.5 \times 10^{-4}$
High	$10 \times 10^{-4}$	$12.5 \times 10^{-4}$
Very high	$20 \times 10^{-4}$	—

The convective fluxes of dry static energy, moisture, cloud water and momentum are calculated in a similar manner, including processes like condensation and precipitation. Besides detrainment, the cloud water  $l_u$  depends on the condensation rate  $c$  of water vapor to cloud water and conversion of cloud water to rain (or autoconversion)  $K$ . The overbar indicates the resolved large-scale variables:

$$\frac{\partial(M_u l_u)}{\partial z} = -D_u l_u + \bar{p}c - \bar{p}l_u K. \quad (2)$$

Here  $\bar{p}$  is the large-scale air density, and  $K$  will be one of the parameters varied for the perturbed physics experiments. The condensation heating includes fusion enthalpy. However, the heat capacity does not consider the liquid or frozen condensate.

Entrainment and detrainment are assumed to consist of an organized and turbulent part:

$$E_u = E_{u,\text{turb}} + E_{u,\text{org}}, \quad (3)$$

$$D_u = D_{u,\text{turb}} + D_{u,\text{org}}. \quad (4)$$

The organized entrainment and detrainment rates are calculated from large-scale and updraft variables and are only applied at certain levels. The turbulent entrainment and detrainment acts on all levels and depends on the mass flux and the entrainment rate parameter  $\epsilon_u$  and detrainment rate parameter  $\delta_u$ :

$$E_{u,\text{turb}} = \epsilon_u M_u, \quad (5)$$

$$D_{u,\text{turb}} = \delta_u M_u. \quad (6)$$

Shallow and deep convection have different entrainment rates, in this study we focus on the entrainment rate for deep (or penetrative) convection. It is assumed that  $\epsilon_u = \delta_u = L^{-1}$  for  $p > p^*$  and  $\epsilon_u = 0$ ,  $\delta_u = L^{-1}$  otherwise, where  $L$  is the characteristic length scale with the standard value  $L = 5$  km, which is also used for our reference experiments, and  $p^*$  is either the arithmetic center of the cloud base and cloud top pressure, or the pressure of maximum updraft velocity (see Möbis and Stevens (2012) for more information). It is usually situated somewhere in the midtroposphere. Therefore the mass flux decreases between  $p^*$  and the level of the cloud top, at which the organized detrainment is applied. In the following, we will refer to both  $\epsilon_u$  and  $\delta_u$  as the turbulent entrainment parameter.

### d. Theoretical moist adiabats

We calculate some theoretical moist adiabats under different assumptions: the pseudoadiabat, which assumes all condensate precipitates immediately, the reversible (or isentropic)

adiabat, which assumes that no condensate is removed, and their respective ice versions that include the fusion enthalpy. We follow the method of Stevens and Siebesma (2020) [section 2b(2)] and start with the enthalpy form of the first law of thermodynamics for an adiabatic process:

$$0 = dh - vdp, \quad (7)$$

where  $dh$  is the change in specific enthalpy,  $v$  is the specific volume and  $dp$  is the change in pressure. The specific enthalpy depends on temperature  $T$ , saturation water vapor mixing ratio  $q_s$ , specific heat capacity  $c_p$  and the phase change enthalpy  $l$ :

$$h = c_p T + lq_s, \quad (8)$$

where  $l$  and  $q_s$  in turn also depend on pressure and temperature and the heat capacity can also change during the ascent. Direct analytical calculation for some types of moist adiabats is possible, but here we take a numerical approach for all. We define

$$dX = dh - vdp \quad (9)$$

and determine  $\partial X/\partial T$  and  $\partial X/\partial p$ . With this we can compute the lapse rate:

$$\frac{dT}{dp} = -\frac{\frac{\partial X}{\partial p}}{\frac{\partial X}{\partial T}}, \quad (10)$$

which can be integrated along  $p$  to obtain a temperature profile. The  $\partial X/\partial T$  and  $\partial X/\partial p$  terms are calculated according to the chosen adiabat. For the pseudoadiabat the specific heat capacity is calculated from the specific heat capacity of dry air and water vapor which decreases with ascent, while for the isentropic adiabat the specific heat capacity of the condensate is also used. The phase change enthalpy  $l$  is simply the condensation enthalpy for the pseudoadiabat and the isentropic adiabat, whereas for their ice-counterparts the fusion enthalpy is added at temperatures below the freezing level. Thus, in the case of freezing we assume that all condensate freezes above the freezing level (and then precipitates in the case of the pseudoice adiabat). The saturation water vapor pressure is calculated with respect to ice in these cases. We start the integration at a certain level in the lower troposphere (usually 700 hPa) where we assume saturation. For the isentropic adiabats we need to specify the total water content in the parcel, for which we chose  $15 \text{ g kg}^{-1}$ . This is larger than the saturation specific humidity at this level since it is likely that the parcel already contains condensate.

In addition, we calculate a lapse rate that considers entrainment using a zero-buoyancy entraining plume model based on the calculation of Singh and O’Gorman (2013) with reference to the pseudoadiabat. In this model, it is assumed that cloud buoyancy is negligible, which in this case means uniform horizontal temperature (virtual temperature effects are not considered). The environmental air is entrained at a rate  $\epsilon = \epsilon_0/z$  and has uniform relative humidity  $r$ , where  $z$  is height above ground. If not stated otherwise, we set  $\epsilon_0 = 0.3$

and  $r = 0.8$ . The vertical profile of the temperature difference to the pseudoadiabat is calculated as

$$\Delta T(z) = -\frac{(1-r)}{1 + \frac{l^2 q_v^*}{R_v T^2 c_p}} \int_{z_b}^z \frac{\epsilon l q_v^*}{c_p} dz', \quad (11)$$

where  $q_v^*$  is the saturation specific humidity of the environment and  $R_v$  is the gas constant for water vapor. The integration is started from the cloud base height  $z_b$  to some height  $z$ . We use 960 hPa as cloud base level (starting the integration at 700 hPa underestimates the effect of entrainment on the temperature profile). In the following, the zero-buoyancy entraining plume will be referred to as “entraining plume.”

### 3. Differences in lapse rates and upper-tropospheric temperatures in conventional climate models

First, we provide an overview of the relationship of lower-tropospheric and upper-tropospheric temperatures in the tropics (defined as  $20^\circ\text{N}$ – $20^\circ\text{S}$ , including both land and ocean). We use the 700 hPa level as proxy to represent the lower free troposphere, since the horizontal temperature variations are small around this level (Bao and Stevens 2021). The WTG approximation holds fairly well and we expect tropospheric temperatures throughout the tropics at and above 700 hPa to be primarily set by deep convection and thus follow a moist adiabatic lapse rate. Consequently, we should be able to infer the upper-tropospheric temperatures from their lower-tropospheric values. In CMIP6 models, this is indeed the case (Fig. 1a), as a model with a warmer lower troposphere tends to have a proportionally warmer upper troposphere and the linear fit across CMIP6 models is at  $1.19 \text{ K K}^{-1}$ . Overall, the models are close to the line that would indicate a pseudoadiabatic relation (moist adiabatic ascent where all condensate precipitates immediately) between 700 and 250 hPa, but the upper-tropospheric temperatures in individual models deviate by up to 1.7 K from the theoretical line. Only very few models are closer to the line that indicates reversible adiabatic relation (convective ascent with no precipitation) than the pseudoadiabatic relation. Despite the robust correlation, there is considerable variability of more than 3 K in the upper-tropospheric temperatures of CMIP6 models for the same lower-tropospheric temperatures. Similar behavior is found in the AMIP simulations (Fig. 1b), which are forced by observational SST datasets (Flannaghan et al. 2014) and, therefore, show less spread in their lower-tropospheric temperatures. However, for a given lower-tropospheric temperature the spread in the upper troposphere has approximately the same magnitude as in the coupled case, resulting in a reduced correlation of lower- and upper-tropospheric temperatures. As in the piControl case, the AMIP experiments can deviate considerably from the temperatures expected from a theoretical pseudoadiabat (up to 1.8 K in the upper troposphere), and even more from the reversible adiabat, while the IUK radiosondes and the ERA5 are fairly close to a pseudoadiabatic relationship (this does not hold for the whole troposphere, as we will show below). The slopes of the regression in both cases in Fig. 1 demonstrate that the cross-model relationship of a proportionally warmer

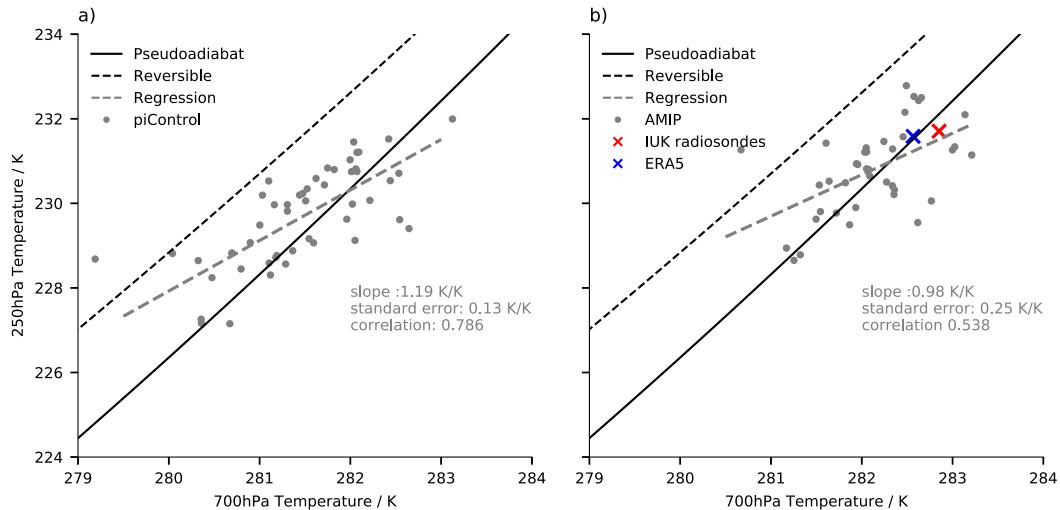


FIG. 1. Tropical ( $20^{\circ}\text{N}$ – $20^{\circ}\text{S}$ ) temperatures in CMIP6 models in the lower and upper troposphere. Shown are (a) preindustrial control simulations and (b) AMIP simulations as well as IUKv2 radiosonde observations (red cross) and ERA5 (blue cross) averaged over years 1979–2014. The black line represents the relationship of 700–250 hPa temperatures under the assumption of a moist pseudoadiabatic ascent and the black dashed line is the same, but for a reversible adiabat. The gray dashed line indicates an ordinary least squares linear regression, with the slope and standard error given in gray text. The cross-model Pearson correlation coefficient is also given.

upper troposphere for models with a warmer lower troposphere does not follow the moist adiabat, as the pseudoadiabatic and the reversible relationships have larger slopes of  $1.99$  and  $1.89\text{ K K}^{-1}$ , respectively. Note that the slopes of the piControl and the AMIP case are within one standard error of each other, while the theoretical adiabats are clearly outside of the standard error range. Figure 1 demonstrates that lower-tropospheric temperatures are useful for predicting upper-tropospheric temperatures, but only to a certain degree. Pseudoadiabatic ascent is a decent approximation for the tropical lapse rate in CMIP6 models but the variation in upper-tropospheric temperatures and the cross-model regression suggests that it is influenced by additional processes that will be discussed in this section.

Above, we have examined the tropical atmosphere in a horizontal mean view. However, only the convective plumes with strong enough buoyancy reach the upper troposphere, which usually originate over warmer SSTs. This coupling of SSTs and convection likely varies across models and has been quantified with precipitation weighted SSTs (PRSST), which show a relationship to upper-tropospheric warming in CMIP5 models (Fueglistaler et al. 2015; Tuel 2019). Here, we apply this methodology to the CMIP6 piControl ensemble using monthly means of SSTs and precipitation to investigate whether this can explain the variations in upper-tropospheric temperatures in the mean state. Additionally, we calculate precipitation weighted 700 hPa temperature (PRTA) in a similar manner, but also include grid points over land in this case, to facilitate the comparison to Fig. 1. We find that PRSSTs cannot explain differences in the mean state (Fig. 2a) any better than plain tropical mean 700 hPa temperatures (Fig. 1a). Instead, the correlation between PRSSTs and upper-tropospheric temperatures is even slightly worse than in Fig. 1a and the spread in upper-tropospheric temperatures also

remains similar at around 4 K for a given value of PRSST. The PRTA (Fig. 2b) shows a better correlation than the PRSST ( $0.75$  vs  $0.68$ ), but worse than the unweighted 700 hPa temperatures ( $0.79$ ). Therefore, the temperatures in the convecting regions, whether SSTs or 700 hPa air temperatures, do not seem to be a better indicator of upper-tropospheric temperatures than the simple tropical mean at 700 hPa. This is likely because the 700 hPa temperatures are homogenized quite effectively to the convecting temperatures by gravity waves (Bao and Stevens 2021) and thus the spatial coupling is naturally included in the tropical mean 700 hPa (or even to some extent in a single radiosonde station). Consequently, we will focus on differences in tropical mean lapse rate behavior above 700 hPa to explain the spread in upper-tropospheric temperatures, given a certain lower-tropospheric temperature.

To illustrate the diversity in tropical lapse rates we assume the moist pseudoadiabat as the closest option to reality (Fig. 1) and show how the tropical lapse rates deviate from the moist pseudoadiabat in individual piControl simulations (AMIP simulations yield similar results). We calculate the pseudoadiabat with the tropical mean 700 hPa temperature as basis and assume saturation at this level. The same calculation was done for radiosonde data from the IUKv2 dataset, as well as the ERA5. CMIP6 models deviate both positively and negatively from their idealized moist pseudoadiabat (Fig. 3). The maximum deviations increase with height, and reach a range from approximately 3 K colder to 2 K warmer than predicted by the moist pseudoadiabat in the upper troposphere. The observations lie within the model spread, but show stronger deviations from the idealized pseudoadiabat than the ensemble mean, especially in the middle troposphere around 500 hPa. Although the reanalysis does not match the observations perfectly, it provides further indication that the real tropical lapse rate is

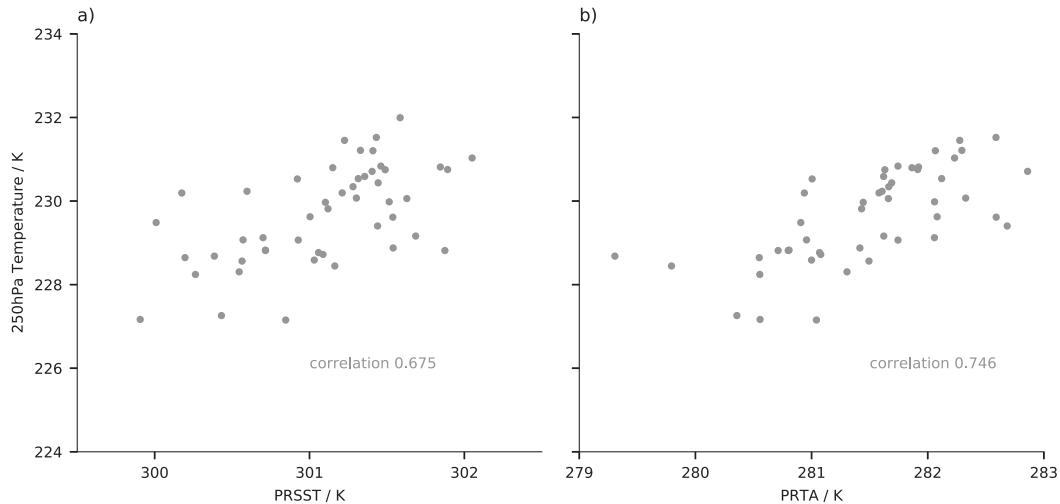


FIG. 2. Tropical ( $20^{\circ}\text{N}$ – $20^{\circ}\text{S}$ ) (a) PRSST and (b) PRTA vs upper-tropospheric temperatures. Both panels show time averages of preindustrial control simulations over the whole available timespan. The cross-model Pearson correlation coefficient is also given.

colder than the pseudoadiabat for most of the troposphere. Models and observations systematically become warmer than their idealized pseudoadiabat above 250 hPa, indicating the transition from the radiative–convective equilibrium to a purely radiative equilibrium (Folkins 2002).

While the pseudoadiabatic ascent seems to reasonably explain the vertical temperature structure in the tropics at first glance, the impact of subgrid-scale processes that alter the diabatic response of the air parcel to the ascent is less clear. In the following we will discuss how some of these processes impact the tropical lapse rate and thereby attempt to explain the spread in lapse rates in CMIP6 models. For example, an air parcel following a reversible adiabat will end up being warmer than one following a pseudoadiabat because of the additional heat reservoir of the condensate contributing to the heat capacity. What happens in reality is somewhere in between these two processes, although it has been suggested that the lapse rate is closer to the pseudoadiabat in the middle and upper troposphere (Bao and Stevens 2021). Figure 3 shows the deviation of the reversible adiabat from the pseudoadiabat (again using the ensemble mean 700 hPa temperature of CMIP6 models) and reveals that the majority of models are closer to the pseudoadiabat. Furthermore, fusion enthalpy causes additional warming during the ascent, which we illustrate with the pseudoice adiabat. Again, the standard pseudoadiabat seems a better fit for most models, but not all. A further process to consider is the entrainment of dry air from outside the cloud that has a cooling effect (Singh and O’Gorman 2013), since it decreases the available moisture and thereby reduces latent heating. The entraining plume approximation, that takes this process into account in a simplified manner (methods), agrees fairly well with a considerable amount of models and the observations. Note that the exact values of the reversible adiabats and the entraining plume are somewhat arbitrary, since they depend on the specified total water content for the parcel and the entrainment rate, respectively.

It is not clear whether the good agreement of the lapse rates to the pseudoadiabat and the entraining plume means that these processes dominate in tropical convection or whether the lapse rates are determined by all of the discussed processes and as a result are close to the moist pseudoadiabat because of compensation. The latter option seems more likely since processes like freezing and (partly) reversible ascent can be observed in reality. In addition, tropospheric temperatures are most likely not set by a single plume with a determined behavior, but rather a spectrum of convecting plumes, that penetrate to different heights and vary in their entrainment rate (Zhou and Xie 2019; Bao and Stevens 2021). We conclude that all of the discussed processes likely impact the lapse rates in CMIP6 models and thereby explain a considerable part of the spread. Since all of these processes happen on subgrid scales, conventional climate models like those in the CMIP6 ensemble parameterize them in a range of different manners. And even if two models use the same convection parameterization, the parameters might be tuned to different values to best compensate errors from other assumptions, which differ across models (Mauritsen et al. 2012). Also, CMIP6 models might contain a common bias related to assumptions made in the convection parameterizations, especially considering the observed lapse rate almost falls outside of the CMIP6 spread at 500 hPa. To demonstrate the impact these subgrid-scale processes have on lapse rates simulated in conventional climate models, we perform experiments with perturbed convection parameterization, which we present in the next section.

#### 4. Processes influencing the tropical lapse rate in ICON-A simulations

We use the atmosphere component ICON-A from the ICON general circulation model in experiments where we perturb the conversion of cloud water to rainfall (auto-conversion) and the turbulent entrainment rate for penetrative convection (the reasoning behind these choices are described

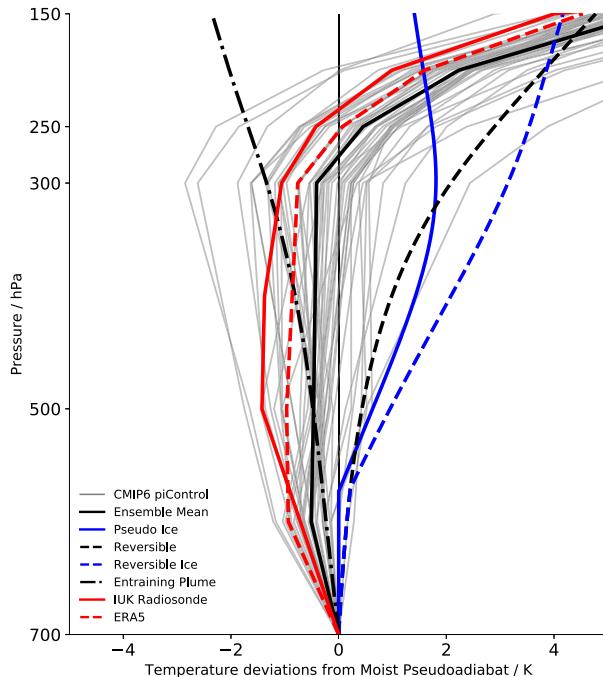


FIG. 3. Deviations from idealized moist pseudoadiabats for all CMIP6 piControl simulations and observations in the tropics ( $20^{\circ}\text{N}$ – $20^{\circ}\text{S}$ ), as well as some idealized cases. First, for every model (and the observations) an idealized moist pseudoadiabat is calculated based on the tropical mean temperature at 700 hPa assuming saturation. The deviations of the actually simulated (and measured) temperatures from each idealized moist pseudoadiabat are illustrated here. In addition, the deviations of the reversible adiabat, freezing pseudoadiabat (assuming fusion above the freezing level), the freezing reversible adiabat, and the entraining plume with respect to the pseudoadiabat are shown.

in the respective subsections below). The values of the tuning parameters are given in Table 1. In the model development process these parameters are set so as to simulate the overall climate as well as possible, hence changing them is generally expected to degrade the climate, at least for those quantities which the tuning process targeted. Nonetheless, by using a skill score for climate models (Reichler and Kim 2008) we demonstrate that all of our experiments but one lie within the skill range of CMIP6 models, and one experiment just outside the range (Fig. 4). The temperature response in both sets of experiments is shown in Fig. 5. Temperatures deviate by more than 2 K from the reference experiment, especially in the upper troposphere (Fig. 5a). With these experiments we can reproduce the spread in temperatures in CMIP6 AMIP simulations (Fig. 5b). The temperature responses in individual experiments and the reasons behind them will be discussed in the following subsections. In this section we present ensemble mean values for our ICON-A experiments, since the differences among ensemble members of one type of experiment are small.

#### a. Conversion of cloud water to rain

First, we examine the experiments with perturbed autoconversion in comparison to the reference experiment. Increasing

autoconversion decreases temperatures in the upper troposphere (Fig. 5) and vice versa. Lower-tropospheric temperatures are unchanged, which means that the lapse rate is different to that of the reference experiment only above 700 hPa. In the case of low autoconversion, the upper troposphere is more than 1 K warmer than the reference experiment.

We analyze the reason behind the temperature changes for the case of low autoconversion: Fig. 6a shows the expected increase of cloud water and cloud ice mixing ratio for decreased autoconversion, as less condensate is precipitated. This increase in condensate points toward two processes that could explain the warming: 1) the condensate contributes to the heat capacity of the air parcel and therefore the expansion cooling is reduced (ascent which is closer to the reversible adiabat than the pseudoadiabat), and 2) the additional cloud water freezes, which produces additional fusion enthalpy. Since the additional condensate heat capacity is not accounted for in the ICON-A convection parameterization, the fusion enthalpy is responsible for the warming. This is supported by the fact that the temperature deviations occur only above 600 hPa level, which is approximately the freezing level (Fig. 5a). We further illustrate the effect of the fusion enthalpy by calculating the temperature deviations from the theoretical pseudoadiabat above 600 hPa (Fig. 6b), with the same method as in Fig. 3. For small values of autoconversion, the lapse rate agrees with the pseudoice adiabat between 600 and 500 hPa before becoming colder, but still remain substantially warmer than the pseudoadiabat. Above 250 hPa, the transition to the purely radiative equilibrium begins, and the idealized moist adiabats become less relevant for understanding the lapse rate. We observe the opposite behavior for increased autoconversion: Cloud condensate decreases, indicating less freezing of condensate cloud water (Fig. 6a), which cools upper-tropospheric temperatures (Fig. 6b).

We conclude that fusion enthalpy can have a considerable impact on upper-tropospheric temperatures and thereby explain some of the spread in CMIP6 models. These experiments demonstrate how the parameterization of autoconversion controls the tropical lapse rate, while also showing a deficiency of the convection parameterization used here, to not consider the effect of the condensate on the heat capacity. In similar experiments with parameterizations that include this effect, even larger temperature deviations can be expected.

#### b. Turbulent entrainment

Changing the entrainment rate for deep convection has substantial effects on the lapse rate and, therefore, upper-tropospheric temperatures (Fig. 5). For increased entrainment we observe a cooling throughout the troposphere with a peak between 500 and 300 hPa. This cooling appears as a shift along the pseudoadiabat at 700 and 250 hPa, (Fig. 5b), but is stronger at around 500 hPa. In contrast, for the case of decreased entrainment, we can observe a warming that is confined to the uppermost troposphere and the tropopause layer (Fig. 5a), and does not change the lapse rate in the mid-troposphere. The mechanisms behind these changes can be illustrated by changes in cloud amount and the heating rates from the convection parameterization (Fig. 7), which will be referred to as convective heating rates.

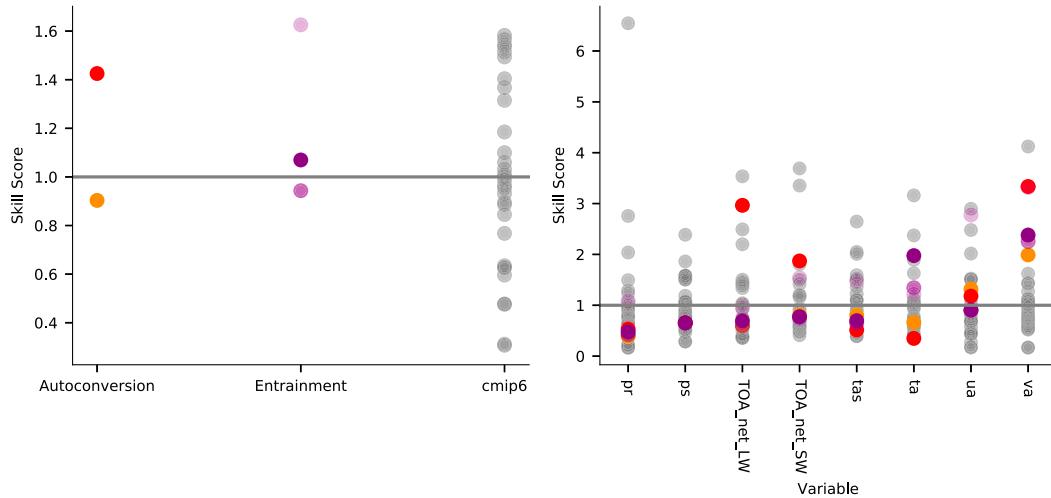


FIG. 4. Skill score of ICON perturbed physics experiments and CMIP6 models in the tropics (30°N–30°S). Gray dots represent CMIP6 AMIP simulations, dots with red shading represent autoconversion experiments, and dots with purple shading represent entrainment experiments. Darker shading indicates stronger autoconversion/entrainment. A skill score of 1 indicates the average CMIP6 AMIP skill score. Variable abbreviations are as follows: pr: precipitation; ps: surface pressure; TOA\_net\_LW: top of the atmosphere net longwave flux; TOA\_net\_SW: top of the atmosphere net shortwave flux; tas: surface temperature; ta: zonal mean temperature; ua: zonal mean zonal wind; and va: zonal mean meridional wind.

For small turbulent entrainment rates, the lower troposphere becomes drier, while the upper troposphere moistens, since less moisture from within the convective plumes is mixed with the surroundings. This is reflected in the cloud fraction changes (Fig. 7a): the lack of entrainment reduces the cloud amount drastically throughout the troposphere and only the anvil cloud amount increases. Weakening entrainment shifts the level of neutral buoyancy higher (Zhou and Xie 2019) since

the ascent is closer to a moist adiabat and, therefore, the convective heating rates (Fig. 7b) increase in the upper troposphere and in the tropopause layer, causing the warming there. Note that at 250 hPa and higher levels, convective heating rates are at least tripled with respect to the reference experiment. Another aspect here could be that reduced entrainment decreases the degree of convective organization (Becker et al. 2017), which also happens in these experiments (not shown)

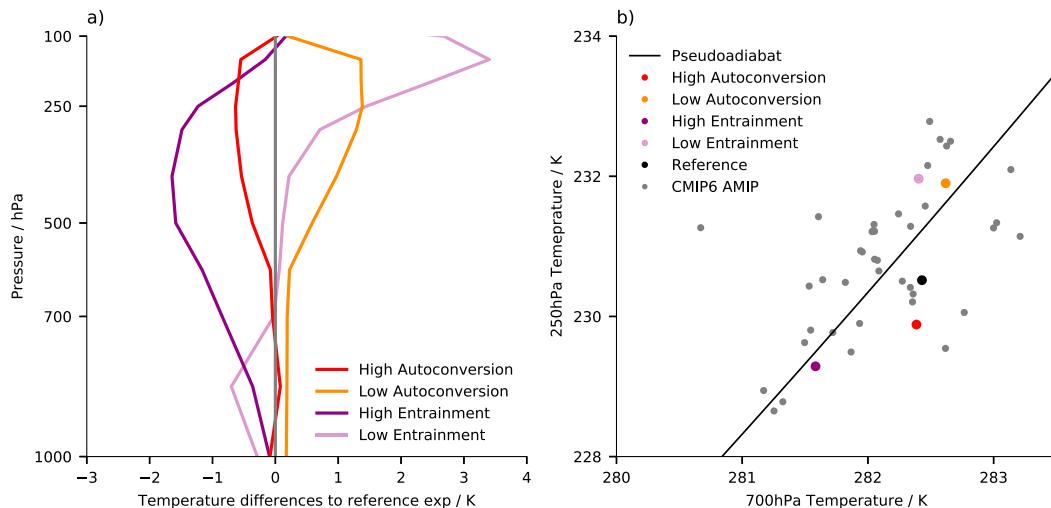


FIG. 5. Changes in atmospheric temperatures in ICON-A perturbed physics experiments. Darker shading represents increased autoconversion (red) and entrainment (purple). (a) Differences in tropical mean (20°N–20°S) temperatures to the reference experiment. (b) Lower- vs upper-tropospheric tropical mean temperatures for all AMIP experiments as well as CMIP6 AMIP simulations. The line represents the relationship expected from a pseudoadiabat.

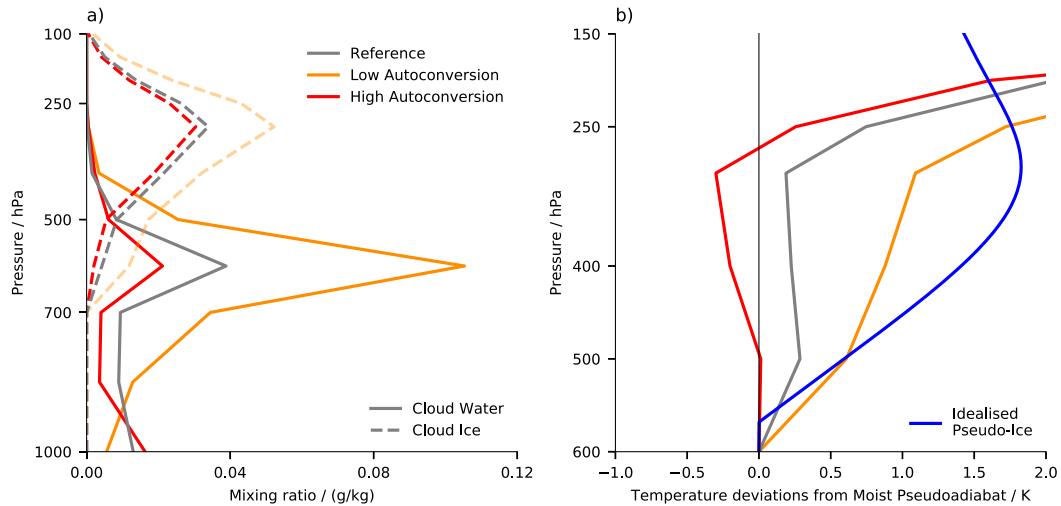


FIG. 6. (a) Cloud water and ice mixing ratio, as well as (b) temperature deviations from the pseudoadiabats in the autoconversion experiments. Darker shading represents increased autoconversion and gray represents the reference experiments. (a) An average over the convecting regions (grid points above the 90th percentile of precipitable water) and (b) the tropical mean ( $20^{\circ}\text{N}$ – $20^{\circ}\text{S}$ ). Deviations from the pseudoadiabats were calculated as in Fig. 3.

and thereby the upper troposphere, where the WTG approximation holds less well (Bao and Stevens 2021), is more uniformly heated by deep convection. Additional fusion enthalpy might also contribute to the warming, since more cloud water is available to freeze in the convective plume, and indeed there is a small positive temperature deviation at around 600 hPa (approximately the freezing level). The cooling in the boundary layer is likely because for reduced entrainment, deep convection can occur at lower temperatures. For large entrainment rates, cloud fraction increases in the lower troposphere and decreases in the upper troposphere, since moisture is detrained to the environment earlier during the ascent. As a result, less condensation heating occurs during the ascent

(Singh and O’Gorman 2013). The reduced convective heating rates are balanced by increased heating rates from the cloud parameterization (not shown), which means that mechanisms outside the convective parameterization, like large scale ascent, control a substantial part of the tropical energy balance in this case. The resulting temperature profile is colder likely because of the reduced latent heating and increased evaporation of detrained cloud water (Mauritsen et al. 2012). This difference in behavior by the parameterizations explains why the vertical temperature response is structured asymmetrically for low and high entrainment rates (Fig. 5a).

We conclude that the entrainment rate also has a substantial impact on tropical lapse rates, demonstrating that the

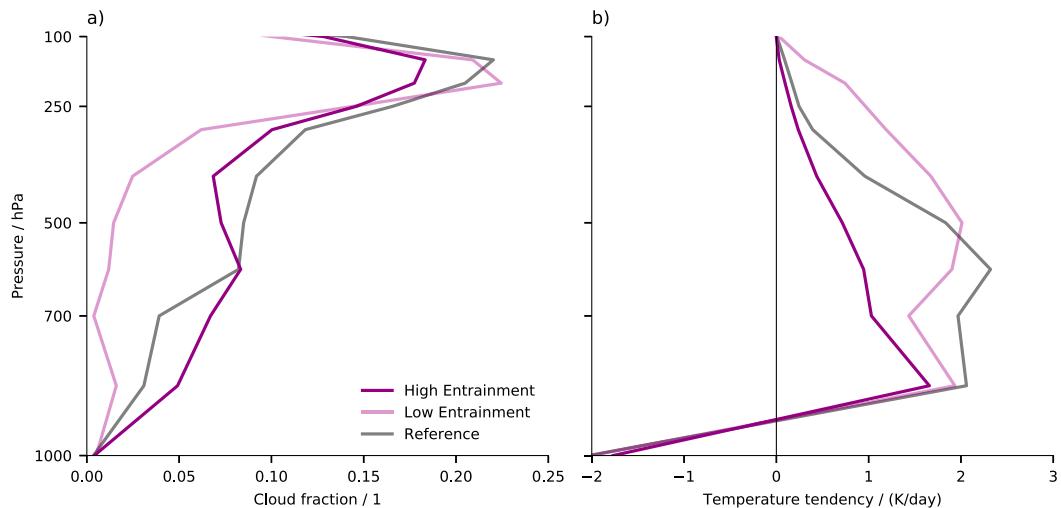


FIG. 7. (a) Tropical ( $20^{\circ}\text{N}$ – $20^{\circ}\text{S}$ ) mean cloud fraction and (b) convective heating rates in the turbulent entrainment experiments. Darker shading represents increased entrainment and gray represents the reference experiments.

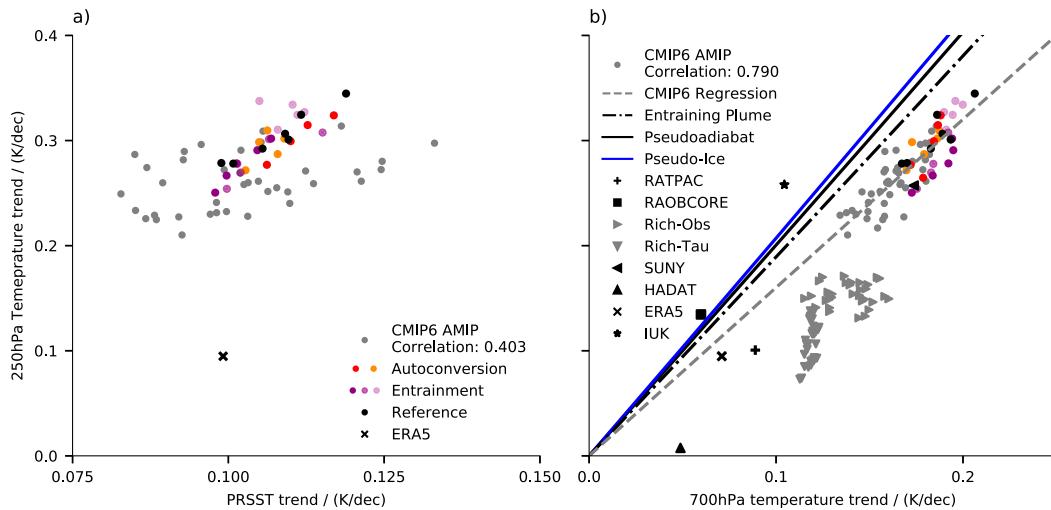


FIG. 8. Tropical ( $20^{\circ}\text{N}$ – $20^{\circ}\text{S}$ ) trends in (a) PRSST and (b) 700 hPa vs trends in 250 hPa temperatures for ICON perturbed physics experiments and CMIP6 AMIP simulations. Darker shading indicates increased autoconversion or entrainment as in Fig. 5. Cross-model correlation is given for CMIP6 models. Trends are calculated over the 1979–2012 period.

representation of entrainment in convection parameterizations is likely one reason behind the spread in upper-tropospheric temperatures in CMIP6 models. The shift of the level of neutral buoyancy illustrates how the levels at which the transition from a convective–radiative equilibrium to a purely radiative equilibrium takes place might be different across models, which has important implications for mid- and upper-tropospheric temperatures. Moreover, increasing entrainment results in temperature anomalies in the middle troposphere that resemble the observed temperature profile (Fig. 3), suggesting that CMIP6 models underestimate the effect of entrainment as suspected by Romps (2010).

### 5. Upper-tropospheric warming

Above, we have demonstrated how the mean tropical lapse rate is influenced by small scale processes like autoconversion and entrainment. Here, we investigate whether these processes could also impact the rate at which the upper troposphere warms under greenhouse gas forcing. In a warmer climate the temperature profile is expected to change to a warmer moist adiabat, thereby amplifying the warming in the upper troposphere with respect to the surface warming. A peak amplification is expected in the upper troposphere around 200 hPa (Santer et al. 2005). However, as for the mean tropical lapse rates investigated in the previous sections, the response to global warming varies across climate models (Santer et al. 2005). More importantly, observations show a substantially weaker warming in the upper troposphere compared to what is simulated by climate models (Santer et al. 2005; Fu et al. 2011; Santer et al. 2017a,b). Here, we assess how coupling of convection and SSTs (Fueglistaler et al. 2015; Tuel 2019; Po-Chedley et al. 2021), and the representation of autoconversion and especially entrainment (Singh and O’Gorman 2013; Zhou and Xie 2019; Miyawaki et al. 2020) impact upper-tropospheric warming in CMIP6 AMIP simulations and our ICON-A experiments. We

show results for the 250 hPa level and the time period of 1979–2012 in this section, because some radiosonde products are only available until 2012. Results are similar for 300 hPa and the 1979–2014 period. There is a considerable influence of internal variability on temperature trends and therefore, unlike in the previous section, we present individual ensemble members for the reference case as well as the autoconversion and entrainment experiments.

Following Fueglistaler et al. (2015) we examine whether the trend in PRSST, which incorporates the coupling of SSTs to convection, explains the different trends in upper-tropospheric warming in CMIP6 models, our ICON-A experiments, and ERA5 (Fig. 8a). We find a large spread in upper-tropospheric temperature and PRSST trends in CMIP6 AMIP simulations but only a weak relationship with a correlation coefficient of 0.403. For our ICON experiments, the relationship is more robust, independent of the parameter experiments. However, the ERA5 upper-tropospheric warming is clearly outside the CMIP6 range, while the PRSST are very similar between ERA5 and CMIP6. Therefore, we conclude that PRSSTs are of limited usefulness in explaining differences in upper-tropospheric warming rates in single realizations of CMIP6 AMIP simulations and cannot explain the difference to the observed upper-tropospheric warming (although the observations also show considerable uncertainty, which will be discussed below). Nevertheless, stronger differences of spatial SST–convection coupling that occur in coupled ocean–atmosphere simulations, are important for explaining even stronger upper-tropospheric warming in these simulations (Tuel 2019; Po-Chedley et al. 2021).

In CMIP6 simulations and our ICON experiments the upper-tropospheric warming is strongly connected to the warming in the lower troposphere (Fig. 8b). As in the case of the mean state (Figs. 1 and 2) lower-tropospheric temperature trends have a stronger correlation to upper-tropospheric temperatures than the PRSST trend has (0.79 vs 0.40). This relationship also seems

to roughly hold for all ICON-A experiments in general and the reference experiments in particular, indicating that variations of trends due to internal variability are vertically consistent throughout the troposphere.

The warming in the upper troposphere shown throughout the models in Fig. 8b is weaker than expected from the theoretical moist adiabats. For stronger entrainment rates ( $\epsilon_0 \approx 0.8$ ) the entraining plume is able to predict the same relationship as the CMIP6 regression suggests (not shown). Most radiosonde estimates show weaker warming trends than the models at both levels, 700 and 250 hPa, with some notable exceptions. The SUNY dataset agrees remarkably well with CMIP6 models and suggests that there is no real discrepancy at all, and the IUK dataset at least shows the same upper-tropospheric warming, but stronger than predicted by the theoretical adiabats. While the other datasets provide gridded data or tropical means, the value for IUK radiosondes shown here is simply the mean over all IUK stations in the tropics (methods). This is a minor issue at 700 hPa (if at all), but apparently produces unrealistically strong upper-tropospheric warming, where the WTG assumption holds less well (Bao and Stevens 2021). In contrast, the HADAT data shows very weak warming, which is also unrealistic, and therefore we will not analyze the IUK and HADAT datasets further. The remaining datasets agree that the upper-tropospheric warming is weaker than simulated by the CMIP6 models, even considering the uncertainties illustrated for the Rich-Obs and Rich-Tau products. However, the radiosondes disagree substantially on the lower-tropospheric warming. For example, the RAOBCORE data suggests that models overestimate the warming in the whole troposphere, suggesting the discrepancy in the upper troposphere is due to the bias at lower altitudes. This contrasts with the Rich-Obs ensemble, according to which CMIP6 models simulate lower-tropospheric warming reasonably well, but the amplification of warming in the upper troposphere is overestimated. It is unclear which radiosonde product is the most trustworthy, but most datasets indicate that the models overestimate recent upper-tropospheric warming. This could be due to a misrepresentation of lower-tropospheric warming, or due to how the warming is amplified from lower to upper troposphere.

Consequently, to quantify the relation of warming in the upper and lower troposphere we calculate an amplification factor, which is the slope of a linear regression of yearly mean, tropical mean temperatures at 700 hPa versus those at 250 hPa. We assess the CMIP6 AMIP ensemble, our ICON-A experiments, observations, the ERA5, and the amplification expected from theoretical adiabats (Fig. 9). For CMIP6 AMIP experiments, lower-tropospheric temperature increases (700 hPa) are amplified by a factor of approximately  $1.5\text{--}1.8 \text{ K K}^{-1}$  in the upper troposphere (250 hPa). The RATPAC, Rich-Obs, Rich-Tau, and SUNY radiosonde estimates fall outside of the CMIP6 range, except for one model, with amplification factors between  $1.35$  and  $1.55 \text{ K K}^{-1}$ . The Raobcore radiosondes show a larger amplification of  $1.6 \text{ K K}^{-1}$ , and amplification in the ERA is even larger, consistent with the CMIP6 ensemble.

While the CMIP6 AMIP ensemble on average overestimates the amplification seen in the radiosonde observations, the prediction by the theoretical adiabats is even stronger and

outside of the model range. For the pseudoadiabat and the pseudoice adiabat the amplification is stronger than  $2 \text{ K K}^{-1}$ . We also show the amplification based on the entraining plume model for a range of entrainment rates from  $\epsilon_0 \approx 0.1$  to  $\epsilon_0 \approx 0.9$ . This illustrates how entrainment tends to decrease the amplification due to the regulation of latent heating.

As indicated before, our ICON-A reference experiments show that the amplification is impacted substantially by natural variability. The spread in the reference experiments demonstrates how internal variability, that is purely driven by the atmosphere in these experiments, covers approximately a third of the CMIP6 spread and is therefore important to consider over the fairly short AMIP observational period. While Suárez-Gutiérrez et al. (2017) have excluded internal variability as sole cause of the lack of upper-tropospheric warming in radiosonde data, Po-Chedley et al. (2021) showed that natural variability can explain a large part of the difference between coupled simulations and satellite data. For our experiments, internal variability alone is unlikely to be the reason for the gap to the weak observed amplification, although the variability is surprisingly strong, considering SSTs are identical across the experiments.

The behavior of the theoretical adiabats is mirrored to some extent by our ICON-A experiments. We can reproduce the spread in the CMIP6 models and close the gap to the observations in our experiments with increased turbulent entrainment rates. The experiments with high entrainment fall just outside the range of the reference experiments. To obtain an even clearer signal, we present experiments with further increased entrainment rates (by a factor of 10, called “very high entrainment,” Table 1), which clearly fall outside of the reference experiment spread and agree well with the Raobcore and SUNY radiosonde datasets, while still producing a climate that is approximately as realistic as the CMIP6 average (Fig. 4). Thus, the entrainment experiments cover almost the complete range of amplification factors in the CMIP6 ensemble and a substantial range of the theoretical entraining plume. The behavior in the autoconversion experiments is less clear. The additional fusion enthalpy should increase amplification slightly, illustrated by the idealized pseudo-ice adiabat. However, the low autoconversion experiments, in which more condensate freezes, show a slightly weaker amplification than the high-autoconversion case (although they both largely fall within the spread of the reference experiments). Overall, changes in autoconversion do not seem to have a substantial effect on the amplification of the historical warming.

The behavior of the entrainment experiments can be understood in a broader context of other recent studies: It has been shown that the entraining plume model gives a more realistic picture of tropical lapse rates and cloud buoyancy for strong warming in idealized radiative-convective equilibrium simulations because it considers reduction in condensation heating through entrainment (Singh and O’Gorman 2013). In addition, Miyawaki et al. (2020) suggest that one reason for weaker warming than predicted by the theoretical moist adiabat in CMIP5 experiments with a quadrupling of  $\text{CO}_2$  is the presence of entrainment, and stronger entrainment weakens upper-tropospheric warming. Our results show that this effect

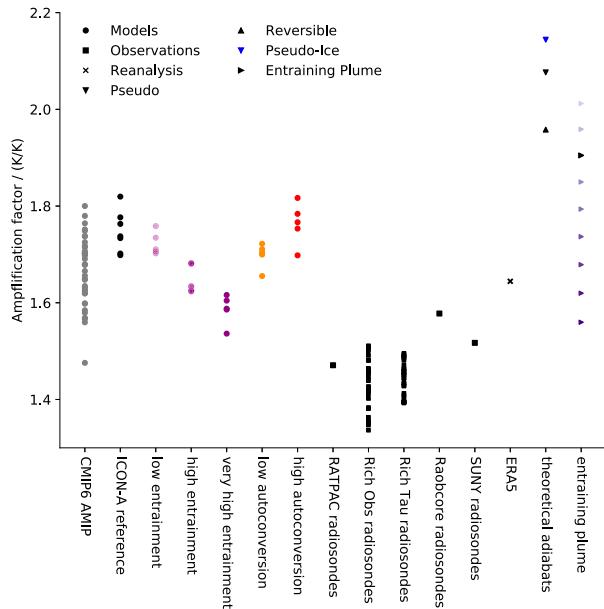


FIG. 9. Amplification factor of tropical mean 250 vs 700 hPa warming for years 1979–2012. Circles show CMIP6 AMIP simulations and ICON perturbed physics experiments, darker color shading indicates stronger entrainment/autoconversion. Black squares show tropical mean radiosonde data. The cross represents the ERA5 and the triangles show the amplification expected from theoretical moist adiabats. In the case of the entraining plume, the amplification is calculated for entrainment rates  $\epsilon_0$  from 0.1 to 0.9 in steps of 0.1.

is already relevant for the historically observed (lack of) warming. Recently, [Zhou and Xie \(2019\)](#) developed a spectral plume model, in which a plume's entrainment rate depends level of neutral buoyancy the plume reaches. This conceptual model reproduces the tropical lapse rate for colder (last glacial maximum) and warmer (RCP 8.5) simulations of general circulation models even more realistically than the single entraining plume. This emphasizes the need to understand the structure of tropical temperatures through convection as a spectrum of entraining plumes with varying characteristics ([Bao and Stevens 2021](#); [Becker and Hohenegger 2021](#)).

## 6. Summary and discussion

Conventional climate models, represented here through the CMIP6 ensemble, show a range of tropical lapse rates and, therefore, models with the same lower-tropospheric temperature can have a spread in their upper-tropospheric temperatures of more than 3 K. Models with similar PRSSTs, that incorporate the spatial coupling of SSTs (or lower-tropospheric temperatures) with convection, show a similar spread in upper-tropospheric temperatures, with a slightly worse correlation than for lower-tropospheric temperatures. Therefore, we focus on explaining the variations in upper-tropospheric temperatures with differences in lapse rates especially between 700 and 250 hPa. While the pseudoadiabatic ascent and the entraining plume provide the best fit to the multimodel mean, individual models can deviate up to 3 K

from the latter in the upper troposphere. Since entrainment, freezing and heat capacity effects all likely play a role, some of these effects could compensate each other to produce a lapse rate close to the pseudoadiabatic. In our ICON perturbed physics experiments we demonstrate how the freezing enthalpy, modified here via the precipitation efficiency or autoconversion, and entrainment substantially alter tropical lapse rates. As these processes are typically used for tuning climate models it suggests that they contribute to the spread in CMIP6 models.

Unlike the mean state, the pseudoadiabatic (or any of the other theoretical adiabats) does not predict the recent upper-tropospheric warming well in models or observations, further indicating that tropical lapse rates are influenced by a multitude of processes. The coupling of convection and SSTs does not explain intermodel differences in upper-tropospheric temperature trends in the AMIP setup, nor the difference to the ERA5 dataset in upper-tropospheric warming. In the CMIP6 ensemble upper-tropospheric warming is closely related to lower-tropospheric warming, but overestimated compared to almost all radiosonde datasets. Furthermore, we can show that most CMIP6 AMIP models fall outside of the observed range of lower to upper-tropospheric amplification, meaning that the upper troposphere warms too strongly for a given lower-tropospheric warming. By increasing entrainment in the convection scheme the amplification falls within the range of observed values, and with varying entrainment rates almost the entire range of amplification factors simulated by CMIP6 models can be covered, which shows that an inadequate representation of entrainment in CMIP6 models likely contributes to the bias between simulated and observed upper-tropospheric warming. However, since the radiosondes disagree among themselves it is not entirely clear whether (i) upper-tropospheric warming is overestimated mainly because the models too strongly amplify a correctly simulated lower-tropospheric warming (as the Rich-Obs dataset indicates), whether (ii) in addition to the overestimation of amplification, the overall warming throughout the troposphere is overestimated ([Mitchell et al. 2013](#)) and the bias already appears at 700 hPa, or (iii) the CMIP6 AMIP ensemble actually simulates lower- and upper-tropospheric warming well, as the SUNY dataset [and to some extent the most recent analysis of satellite data, [Po-Chedley et al. \(2021\)](#)] suggests. If the warming is indeed already too weak in the lower troposphere, the question remains what causes this bias. In a nutshell, according to most radiosonde products, CMIP6 models overestimate upper-tropospheric warming, and according to all radiosonde products on average overestimate the amplification from lower to upper troposphere, which depends on the representation of convective entrainment.

Our results suggest that various small scale processes that occur in moist convection impact tropical lapse rates, but the exact contribution of each process remains unclear. Consequently, climate models not capturing the weak upper-tropospheric warming is no surprise considering our lack of understanding in what determines the tropical lapse rate. With the results presented here, that emphasize the important role of entrainment, we connect the theoretical, more idealized work by [Singh and O'Gorman \(2013\)](#) and

Zhou and Xie (2019) to studies analyzing the weak observed warming (Santer et al. 2005; Fu et al. 2011; Santer et al. 2017a,b) and show that entrainment dampens the recent warming in a realistic AMIP setup. While comparing turbulent entrainment used here and in other parameterizations to real observed entrainment is difficult, it has been suggested that entrainment in models is often underestimated (Romps 2010). We can only speculate that CMIP6 models have not been tuned with higher entrainment rates, because that leads to a worse representation of the overall climate (the energy balance in the tropics changes substantially in our high entrainment experiments; see section 4b). A more realistic representation of convection and entrainment could be achieved by applying a more sophisticated spectral cumulus parameterization, in which the entrainment rate depends cloud characteristics (Zhou and Xie 2019), as in Baba and Giorgetta (2019). The deficiencies of convection parameterizations are also highlighted in the autoconversion experiments, in which the variation of cloud water and cloud ice should affect the heat capacity and therefore temperatures, but this process is simply not included in the parameterization used in the ICON-A model.

While the relationship between PRSSTs and upper-tropospheric warming is seemingly weak in our results, unlike for Fueglistaler et al. (2015), we only present individual ensemble members, and use CMIP6 instead of CMIP5. The relationship for individual ensemble members is more robust in CMIP5, likely because CMIP5 uses two different SST datasets which increases the spread in PRSST [see Fig. 1 in Fueglistaler et al. (2015)]. Also, we omit any analysis of ensemble means, and therefore would argue that our results are not directly comparable to those of Fueglistaler et al. (2015). Especially for stronger variations in PRSSTs that occur in coupled atmosphere–ocean models, the coupling of convection and SSTs is an important factor to explain differences in upper-tropospheric warming among models (Mitchell et al. 2013; Tuel 2019; Po-Chedley et al. 2021). CMIP6 AMIP models seem to simulate a realistic trend in PRSSTs, which should be more reliably observed than upper-tropospheric warming. Therefore, if there is a bias in upper-tropospheric warming in CMIP6 AMIP models, as most radiosonde products suggest, then this bias originates from other misrepresented processes, of which entrainment is likely one.

A large part of the presented analysis was also done for CMIP5 models, but overall did not differ much from the CMIP6 results presented here. While CMIP5 models include some stronger outliers for tropical mean lapse rates, the spread in the amplification factor is virtually identical, suggesting that conventional climate models have not significantly improved in this regard over the past decade, even though our understanding of the problem with regards to the role of entrainment (Singh and O’Gorman 2013; Miyawaki et al. 2020), and the coupling of SSTs and convection (Fueglistaler et al. 2015; Tuel 2019; Po-Chedley et al. 2021) has increased and errors in satellite as well as radiosonde observations have reduced. Therefore, we propose to tackle this problem in the future with high-resolution storm resolving models, that do not rely on parameterized convection (Stevens et al. 2019). While cloud microphysical properties will

always have to be parameterized, at least an improvement in the representation of entrainment and overall tropical circulation can be expected (Stevens et al. 2020). Nevertheless, long term or large ensemble simulations will likely rely on convection parameterizations for the near future, and therefore further development in this area (Baba and Giorgetta 2019) should be beneficial for the representation of global circulation in climate models.

*Acknowledgments.* The research is supported by public funding to the Max Planck Society. The authors also acknowledge support from the European Union’s Horizon 2020 Research and Innovation Programme under Grant 820829 (CONSTRAIN Project). The authors thank C. Wengel, C. Hohenegger, T. Dauhut, and L. Kluff for valuable discussion and feedback.

*Data availability statement.* Run scripts for the ICON experiments and code to produce the figures are available here: [https://github.com/pkeil7/tropical\\_lapse\\_rates](https://github.com/pkeil7/tropical_lapse_rates). CMIP6 data are available at DKRZ or from the Earth System Grid Federation (ESGF) (<https://esgf-node.llnl.gov/projects/cmip6/>). Rich and RAOBCORE radiosonde data are available here: <https://www.univie.ac.at/theoret-met/research/raobcore/>. SUNY-Albany radiosonde dataset is available here: <ftp://aspen.atmos.albany.edu/data/UA-HRD/>. RATPAC data are available here: <https://www.ncdc.noaa.gov/data-access/weather-balloon/radiosonde-atmospheric-temperature-products-accessing-climate>. IUK data are available here: <https://www.ccr.unsw.edu.au/professor-steven-sherwood/research-steven-sherwood/iuk-radiosonde-analysis-project-now-updated>. ERA5 data are available on the Copernicus Climate Change Service Climate Data Store (CDS): <https://cds.climate.copernicus.eu/cdsapp#!/home>.

## REFERENCES

- Baba, Y., and M. A. Giorgetta, 2019: Tropical variability simulated in ICON-A with a spectral cumulus parameterization. *J. Adv. Model. Earth Syst.*, **12**, e2019MS001732, <https://doi.org/10.1029/2019MS001732>.
- Bao, J., and B. Stevens, 2021: The elements of the thermodynamic structure of the tropical atmosphere. *Earth Space Sci.*, **65**, <https://doi.org/10.1002/essoar.10507203.1>.
- Becker, T., and C. Hohenegger, 2021: Entrainment and its dependency on environmental conditions and convective organization in convection-permitting simulations. *Mon. Wea. Rev.*, **149**, 537–550, <https://doi.org/10.1175/MWR-D-20-0229.1>.
- , B. Stevens, and C. Hohenegger, 2017: Imprint of the convective parameterization and sea-surface temperature on large-scale convective self-aggregation. *J. Adv. Model. Earth Syst.*, **9**, 1488–1505, <https://doi.org/10.1002/2016MS000865>.
- Bretherton, C. S., and P. K. Smolarkiewicz, 1989: Gravity waves, compensating subsidence and detrainment around cumulus clouds. *J. Atmos. Sci.*, **46**, 740–759, [https://doi.org/10.1175/1520-0469\(1989\)046<0740:GWCSAD>2.0.CO;2](https://doi.org/10.1175/1520-0469(1989)046<0740:GWCSAD>2.0.CO;2).
- Crueger, T., and Coauthors, 2018: ICON-A, the atmosphere component of the ICON earth system model: II. Model evaluation. *J. Adv. Model. Earth Syst.*, **10**, 1638–1662, <https://doi.org/10.1029/2017MS001233>.
- Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, 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>.
- Flannaghan, T. J., S. Fueglistaler, I. M. Held, S. Po-Chedley, B. Wyman, and M. Zhao, 2014: Tropical temperature trends in atmospheric general circulation model simulations and the impact of uncertainties in observed SSTs. *J. Geophys. Res. Atmos.*, **119**, 13 327–13 337, <https://doi.org/10.1002/2014JD022365>.
- Folkins, I., 2002: Origin of lapse rate changes in the upper tropical troposphere. *J. Atmos. Sci.*, **59**, 992–1005, [https://doi.org/10.1175/1520-0469\(2002\)059<0992:OOLRCI>2.0.CO;2](https://doi.org/10.1175/1520-0469(2002)059<0992:OOLRCI>2.0.CO;2).
- Free, M., D. J. Seidel, J. K. Angell, J. Lanzante, I. Durre, and T. C. Peterson, 2005: Radiosonde atmospheric temperature products for assessing climate (RATPAC): A new data set of large-area anomaly time series. *J. Geophys. Res.*, **110**, D22101, <https://doi.org/10.1029/2005JD006169>.
- Fu, Q., S. Manabe, and C. M. Johanson, 2011: On the warming in the tropical upper troposphere: Models versus observations. *Geophys. Res. Lett.*, **38**, L15704, <https://doi.org/10.1029/2011GL048101>.
- Fueglistaler, S., C. Radley, and I. M. Held, 2015: The distribution of precipitation and the spread in tropical upper tropospheric temperature trends in CMIP5/AMIP simulations. *Geophys. Res. Lett.*, **42**, 6000–6007, <https://doi.org/10.1002/2015GL064966>.
- Giorgetta, M. A., and Coauthors, 2018: ICON-A, the atmosphere component of the ICON earth system model: I. Model description. *J. Adv. Model. Earth Syst.*, **10**, 1613–1637, <https://doi.org/10.1029/2017MS001242>.
- Haimberger, L., C. Tavalato, and S. Sperka, 2012: Homogenization of the global radiosonde temperature dataset through combined comparison with reanalysis background series and neighboring stations. *J. Climate*, **25**, 8108–8131, <https://doi.org/10.1175/JCLI-D-11-00668.1>.
- Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999–2049, <https://doi.org/10.1002/qj.3803>.
- Lee, S., C. Woods, and R. Caballero, 2019: Relation between Arctic moisture flux and tropical temperature biases in CMIP5 simulations and its fingerprint in RCP8.5 projections. *Geophys. Res. Lett.*, **46**, 1088–1096, <https://doi.org/10.1029/2018GL080562>.
- Lu, J., and M. Cai, 2010: Quantifying contributions to polar warming amplification in an idealized coupled general circulation model. *Climate Dyn.*, **34**, 669–687, <https://doi.org/10.1007/s00382-009-0673-x>.
- Mauritsen, T., and Coauthors, 2012: Tuning the climate of a global model. *J. Adv. Model. Earth Syst.*, **4**, M00A01, <https://doi.org/10.1029/2012MS000154>.
- Mitchell, D. M., P. W. Thorne, P. A. Stott, and L. J. Gray, 2013: Revisiting the controversial issue of tropical tropospheric temperature trends. *Geophys. Res. Lett.*, **40**, 2801–2806, <https://doi.org/10.1002/grl.50465>.
- Miyawaki, O., Z. Tan, T. A. Shaw, and M. F. Jansen, 2020: Quantifying key mechanisms that contribute to the deviation of the tropical warming profile from a moist adiabat. *Geophys. Res. Lett.*, **47**, e2020GL089136, <https://doi.org/10.1029/2020GL089136>.
- Möbis, B., and B. Stevens, 2012: Factors controlling the position of the intertropical convergence zone on an aquaplanet. *J. Adv. Model. Earth Syst.*, **4**, M00A04, <https://doi.org/10.1029/2012MS000199>.
- Nordeng, T. E., 1994: Extended versions of the convective parameterization scheme at ECMWF and their impact on the mean and transient activity of the model in the tropics. ECMWF Tech. Memo. 206, 42 pp., <https://doi.org/10.21957/e34xwhysw>.
- Plant, R. S., and J.-I. Yano, 2016: *Parameterization of Atmospheric Convection*. World Scientific Press, 1172 pp.
- Po-Chedley, S., T. J. Thorsen, and Q. Fu, 2015: Removing diurnal cycle contamination in satellite-derived tropospheric temperatures: Understanding tropical tropospheric trend discrepancies. *J. Climate*, **28**, 2274–2290, <https://doi.org/10.1175/JCLI-D-13-00767.1>.
- , B. D. Santer, S. Fueglistaler, M. D. Zelinka, P. J. Cameron-Smith, J. F. Painter, and Q. Fu, 2021: Natural variability contributes to model–satellite differences in tropical tropospheric warming. *Proc. Natl. Acad. Sci. USA*, **118**, e2020962118, <https://doi.org/10.1073/pnas.2020962118>.
- Reichler, T., and J. Kim, 2008: How well do coupled models simulate today's climate? *Bull. Amer. Meteor. Soc.*, **89**, 303–311, <https://doi.org/10.1175/BAMS-89-3-303>.
- Romps, D. M., 2010: A direct measure of entrainment. *J. Atmos. Sci.*, **67**, 1908–1927, <https://doi.org/10.1175/2010JAS3371.1>.
- , and Z. Kuang, 2010: Do undiluted convective plumes exist in the upper tropical troposphere? *J. Atmos. Sci.*, **67**, 468–484, <https://doi.org/10.1175/2009JAS3184.1>.
- Santer, B. D., and Coauthors, 2005: Atmospheric science: Amplification of surface temperature trends and variability in the tropical atmosphere. *Science*, **309**, 1551–1556, <https://doi.org/10.1126/science.1114867>.
- , and Coauthors, 2017a: Causes of differences in model and satellite tropospheric warming rates. *Nat. Geosci.*, **10**, 478–485, <https://doi.org/10.1038/ngeo2973>.
- , and Coauthors, 2017b: Comparing tropospheric warming in climate models and satellite data. *J. Climate*, **30**, 373–392, <https://doi.org/10.1175/JCLI-D-16-0333.1>.
- Sherwood, S. C., and N. Nishant, 2015: Atmospheric changes through 2012 as shown by iteratively homogenized radiosonde temperature and wind data (IUKv2). *Environ. Res. Lett.*, **10**, 054007, <https://doi.org/10.1088/1748-9326/10/5/054007>.
- , J. R. Lanzante, and C. L. Meyer, 2005: Radiosonde daytime biases and late-20th century warming. *Science*, **309**, 1556–1559, <https://doi.org/10.1126/science.1115640>.
- Simmons, A., and Coauthors, 2020: Global stratospheric temperature bias and other stratospheric aspects of ERA5 and ERA5.1. Tech. Memo. 859, 38 pp., <https://doi.org/10.21957/rcxqfmg0>.
- Singh, M. S., and P. A. O’Gorman, 2013: Influence of entrainment on the thermal stratification in simulations of radiative-convective equilibrium. *Geophys. Res. Lett.*, **40**, 4398–4403, <https://doi.org/10.1002/grl.50796>.
- Sohn, B. J., S. Lee, E. S. Chung, and H. J. Song, 2016: The role of the dry static stability for the recent change in the Pacific Walker circulation. *J. Climate*, **29**, 2765–2779, <https://doi.org/10.1175/JCLI-D-15-0374.1>.
- Stevens, B., and A. P. Siebesma, 2020: Clouds as fluids. *Clouds and Climate: Climate Science’s Greatest Challenge*, A. Pier Siebesma et al., Eds., Cambridge University Press, 35–73, <https://doi.org/10.1017/9781107447738.003>.
- , and Coauthors, 2019: DYAMOND: The DYNAMics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains. *Prog. Earth Planet. Sci.*, **6**, 61, <https://doi.org/10.1186/s40645-019-0304-z>.
- , and Coauthors, 2020: The added value of large-eddy and storm-resolving models for simulating clouds and precipitation. *J. Meteor. Soc. Japan*, **98**, 395–435, <https://doi.org/10.2151/jmsj.2020-021>.

- Stone, P. H., and J. H. Carlson, 1979: Atmospheric lapse rate regimes and their parameterization. *J. Atmos. Sci.*, **36**, 415–423, [https://doi.org/10.1175/1520-0469\(1979\)036<0415:ALRRAT>2.0.CO;2](https://doi.org/10.1175/1520-0469(1979)036<0415:ALRRAT>2.0.CO;2).
- Suárez-Gutiérrez, L., C. Li, P. W. Thorne, and J. Marotzke, 2017: Internal variability in simulated and observed tropical tropospheric temperature trends. *Geophys. Res. Lett.*, **44**, 5709–5719, <https://doi.org/10.1002/2017GL073798>.
- Thorne, P. W., D. E. Parker, S. F. Tett, P. D. Jones, M. McCarthy, H. Coleman, and P. Brohan, 2005: Revisiting radiosonde upper air temperatures from 1958 to 2002. *J. Geophys. Res. Atmos.*, **110**, D18105, <https://doi.org/10.1029/2004JD005753>.
- , and Coauthors, 2007: Tropical vertical temperature trends: A real discrepancy? *Geophys. Res. Lett.*, **34**, L16702, <https://doi.org/10.1029/2007GL029875>.
- , J. R. Lanzante, T. C. Peterson, D. J. Seidel, and K. P. Shine, 2011: Tropospheric temperature trends: History of an ongoing controversy. *Wiley Interdiscip. Rev.: Climate Change*, **2**, 66–88, <https://doi.org/10.1002/wcc.80>.
- Tiedtke, M., 1989: A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Mon. Wea. Rev.*, **117**, 1779–1800, [https://doi.org/10.1175/1520-0493\(1989\)117<1779:ACMFSF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2).
- Trabing, B. C., M. M. Bell, and B. R. Brown, 2019: Impacts of radiation and upper-tropospheric temperatures on tropical cyclone structure and intensity. *J. Atmos. Sci.*, **76**, 135–153, <https://doi.org/10.1175/JAS-D-18-0165.1>.
- Tuel, A., 2019: Explaining differences between recent model and satellite tropospheric warming rates with tropical SSTs. *Geophys. Res. Lett.*, **46**, 9023–9030, <https://doi.org/10.1029/2019GL083994>.
- Wu, Y., M. Ting, R. Seager, H. P. Huang, and M. A. Cane, 2011: Changes in storm tracks and energy transports in a warmer climate simulated by the GFDL CM2.1 model. *Climate Dyn.*, **37**, 53–72, <https://doi.org/10.1007/s00382-010-0776-4>.
- Zelinka, M. D., and D. L. Hartmann, 2010: Why is longwave cloud feedback positive? *J. Geophys. Res.*, **115**, D16117, <https://doi.org/10.1029/2004JD005753>.
- Zhou, C., J. Wang, A. Dai, and P. W. Thorne, 2021: A new approach to homogenize global subdaily radiosonde temperature data from 1958 to 2018. *J. Climate*, **34**, 1163–1183, <https://doi.org/10.1175/JCLI-D-20-0352.1>.
- Zhou, W., and S. P. Xie, 2019: A conceptual spectral plume model for understanding tropical temperature profile and convective updraft velocities. *J. Atmos. Sci.*, **76**, 2801–2814, <https://doi.org/10.1175/JAS-D-18-0330.1>.