Tropical free-tropospheric humidity differences and their effect on the clear-sky radiation budget in global strom-resolving models

Theresa Lang[1,2], Ann Kristin Naumann [3,1], Bjorn Stevens [3], Stefan A. Buehler [1]

[1]Meteorological Institute, Center for Earth System Research and Sustainability (CEN), Universität Hamburg, Hamburg, Germany
[2]International Max Planck Research School on Earth System Modelling, Max Planck Institute for Meteorology, Hamburg, Germany
[3]Max Planck Institute for Meteorology, Hamburg, Germany

Key Points:

- In global storm-resolving models the spread in free-tropospheric humidity is approximately halved compared to conventional climate models
- The remaining humidity differences still cause a considerable spread of 1.2 $\mathrm{W m^{-2}}$ in tropical mean clear-sky OLR
- Reducing humidity biases would be most beneficial in the lower and mid free troposphere, particularly in dry subsidence regimes and close to deep convective regimes

Corresponding author: Theresa Lang, `theresa.lang@uni-hamburg.de`

Abstract

Although the humidity of the tropical free-troposphere plays a key role in controlling the Earth's energy budget, it is poorly simulated by conventional climate models. Recently developed global storm-resolving models (GSRMs) are expected to better represent the relevant processes, but it is unclear to what extent humidity biases are reduced. In this study we quantify inter-model differences in tropical free-tropospheric humidity and their impact on the clear-sky radiation budget in an ensemble of nine GSRMs called DYAMOND. We find that throughout most of the free troposphere the inter-model spread in relative humidity (RH) is approximately halved compared to conventional climate models. Nevertheless, the remaining differences cause a considerable spread of $1.2\,\mathrm{Wm}^{-2}$ in tropical mean clear-sky outgoing longwave radiation (OLR). This spread is mainly caused by RH differences in the lower and mid free troposphere, whereas RH differences in the upper troposphere have a minor impact. By examining model differences in moisture space we identify two regimes with a particularly large contribution to the spread in tropical mean OLR: rather moist regions at the transition from deep convective to subsidence regimes and very dry subsidence regimes. In the regions identified as most critical we do not find a direct relation between the RH differences and differences in the RH transport by the resolved circulation, suggesting that inter-model differences are mainly related to unresolved processes like microphysics and turbulence. Hence, a more detailed understanding of how these processes affect RH is needed to further constrain the humidity distribution in GSRMs.

Plain Language Summary

The humidity of the atmosphere affects radiation and hence the Earth's energy budget, but it is poorly simulated by conventional climate models. In this study we investigate whether recently developed high-resolution models simulate humidity more accurately. We find that humidity biases in the tropics are approximately halved compared to conventional climate models. Nevertheless, the humidity biases still have a considerable effect on the radiation budget. We also investigate in which regions of the tropics a further reduction of biases would be most beneficial. In the vertical, it is the altitude region between about 1 km and 10 km. In the horizontal, we find two tropical regimes that are particularly important: Dry regimes with very strong subsidence and moister regimes at the edge of deep convective regimes. In the regions we identify as most important the humidity biases are most likely related to processes that are still unresolved in high-resolution models. Therefore, a better understanding of how these processes affect humidity is needed.

1 Introduction

The humidity distribution in the tropical free troposphere plays an important role in controlling the Earth's radiation budget. However, it is poorly simulated by conventional climate models (e.g. Jiang et al., 2012). A major uncertainty arises from parameterizations of unresolved processes, particularly the convective parameterization. Global storm-resolving models (GSRMs) forgo the parameterization of deep convection for the first time on a global scale and therefore raise hopes for an improvement of the simulated humidity distribution, but it is unclear to what extent these hopes are justified. In this study we quantify differences in the distribution of tropical free-tropospheric humidity as well as the resulting spread in clear-sky outgoing longwave radiation (OLR) in GSRMs based on the DYAMOND multi-model ensemble. Moreover, we identify the regions of the tropical atmosphere, which would most benefit from a reduction in humidity biases.

Water vapour is the most important absorber of infrared radiation in the atmosphere and hence strongly impacts the Earth's OLR (e.g. Harries, 1997). Furthermore, it is associated with a feedback that amplifies the climate system's response to forcings such as an increase in anthropogenic greenhouse gasses (e.g. Held & Soden, 2000). A region that has received particular attention in this context is the free troposphere, because OLR is particularly sensitive to humidity changes there (Spencer & Braswell, 1997; Held & Soden, 2000; Soden et al., 2005). Hence, for the energy budget it is crucial that the distribution of free-tropospheric humidity is well reproduced by climate models.

However, there are substantial errors in the simulation of the present-day distribution of free-tropospheric humidity by General Circulation Models (GCMs). Several studies revealed significant inter-model spreads in ensembles of GCMs, both in models forced by observed sea surface temperatures (e.g. Brogniez et al., 2005) and fully coupled atmosphere ocean models (e.g. Pierce et al., 2006; John & Soden, 2007; Jiang et al., 2012). These studies also consistently found a moist bias in the tropical free troposphere with respect to satellite observations. To date it is not clear which physical processes control the inter-model differences and biases.

Processes affecting the tropical free-tropospheric humidity distribution act on a variety of scales, not all of which are well represented in GCMs (Sherwood et al., 2010). While transport of humidity by the large-scale circulation is explicitly resolved, sub-grid-scale processes, like convective and turbulent mixing as well as microphysical processes are only crudely represented in the form of parameterizations. The relative importance of these processes in setting the free-tropospheric humidity distribution has been studied extensively. Diagnostic studies indicate that the observed humidity distribution can be reproduced reasonably well only considering advection of convectively-saturated air on scales resolved by GCMs (e.g. Sherwood, 1996; Pierrehumbert & Roca, 1998; Dessler & Sherwood, 2000). This suggests that after the air is saturated in deep convection, sources and sinks of water vapour from phase changes or mixing are modest. On the other hand, since these are the most poorly constrained processes in the models, it is likely that differences in their parameterizations play a major role in controlling inter-model differences in the humidity distribution.

A promising step towards reducing the uncertainty in the humidity distribution has been made with the development of GSRMs (Satoh et al., 2019). These models solve the non-hydrostatic equations on global grids with kilometre-scale resolution. At such resolutions the models begin to resolve precipitating convective systems and therefore forgo the need to parameterize deep convection. It is hoped that this eradicates some long-standing biases associated with this parameterization (e.g. Stevens & Bony, 2013; Stevens et al., 2020). It certainly also eliminates an important source of uncertainty for the distribution of free-tropospheric humidity. However, uncertainty remains since deep convection is imperfectly resolved at kilometre-scale resolution (e.g. Bryan et al., 2003; Miyamoto et al., 2013), and because other relevant processes like shallow convection, turbulent mixing and microphysical processes remain unresolved or poorly resolved in GSRMs.

In this study we investigate whether and by how much the spread in free-tropospheric humidity is reduced in GSRMs compared to GCMs. To do so, we quantify the spread in a multi-model ensemble consisting of nine GSRMs, which took part in a first intercomparison called DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains (DYAMOND) (Stevens et al., 2019). As a first step towards narrowing down the processes responsible for the remaining humidity differences, we investigate whether they are related to differences in the resolved humidity transport.

From the magnitude of the humidity differences alone it is hard to assess how relevant they are for the radiation budget, since the sensitivity of OLR to a given humidity perturbation varies both with the altitude at which the perturbation is applied and with the humidity of the base state (e.g. Spencer & Braswell, 1997). Therefore, we translate the humidity differences into differences in clear-sky OLR using a radiative transfer model. Furthermore, we identify those regions in the tropical atmosphere, in which a future reduction of humidity differences is most effective in reducing differences in clear-sky OLR.

We perform the comparison of the DYAMOND models in moisture space, i.e. we sort the atmospheric state from dry to moist. This allows us, on the one hand, to distinguish between different dynamic regimes of the tropics, which is useful for identifying the sources of inter-model differences as well as for understanding differences in OLR. On the other hand, humidity fields in moisture space are highly aggregated, which ensures robust statistics. The representation of the atmosphere in moisture space is inspired by Bretherton et al. (2005), who used it to study the energy balance of convective self-aggregation in radiative-convective equilibrium simulations. Later, the depiction in moisture space has also proven useful for analysing observational data (Schulz & Stevens, 2018) and to bypass the issue of co-location when comparing observations and model simulations (Naumann & Kiemle, 2020).

This paper is organized as follows: In Section 2 we introduce the DYAMOND simulations and describe our post-processing of the model output. In Section 3 we quantify inter-model humidity differences in the tropical mean and in moisture space. Moreover, we investigate whether humidity anomalies in the models are related to anomalies in the humidity transport by the resolved circulation. The impact of the humidity differences on the clear-sky radiation budget is examined in Section 4.

2 DYAMOND simulations

2.1 Models and experimental protocol

DYAMOND is the first intercomparison project for GSRMs, comparing 40-day simulations of nine models (only acronyms are given here): ICON, NICAM, ARPEGE-NH, FV3, GEOS, MPAS, UM, SAM and IFS. In the following we provide a brief overview of the models and the experimental protocol of DYAMOND. A more detailed description is given by Stevens et al. (2019).

Most of the DYAMOND models solve the fully compressible non-hydrostatic Navier-Stokes equations. Two exceptions are SAM, which uses the anelastic form of the non-hydrostatic equations, and IFS, which solves the primitive equations and is hence a hydrostatic model. The models solve their governing equations on a variety of different numerical grids. The horizontal grid spacing is between 2.5 km and 5 km in eight of the nine models. The only exception is UM, which uses a latitude-longitude grid with a somewhat coarser resolution at low latitudes (7.8 km at the equator). The number of vertical levels and the vertical extent of the model grid also vary between the models. None of the models are tuned at such high resolution.

The models also differ in the parameterizations used to represent unresolved processes. In particular, there are different approaches to handle convection, reflecting some disagreement about which motions are adequately resolved at kilometre-resolution. While in some models convection is not parameterized at all, in others shallow convection is parameterized. GEOS and MPAS even employ scale-aware parameterizations for deep convection. There is also diversity in the parameterizations for boundary layer turbulence and microphysics.

The DYAMOND simulations were run for 40 days from 1 August to 10 September 2016. They were initialized with common atmospheric fields from the ECMWF global (9 km) meteorological analysis. Daily sea surface temperatures (SSTs) and sea ice concentrations from the ECMWF analysis were used as boundary conditions. The initialization of the land surface was left to the practices of the individual modelling groups. After the initialization each simulation was allowed to evolve freely without further forcing.

2.2 Post-processing and profile selection

We use the 3-hourly output of atmospheric pressure p, temperature T, specific humidity q as well as the three components of the wind field U, V and W. Following Stevens et al. (2019) we exclude the first ten days of the simulations and only use the last 30 days to minimize the effects of biases from differences in the model spin-up as well as constraints from the common initialization. For each model the fields are horizontally interpolated from the native model grid to a common regular latitude-longitude grid covering the tropics (30° S to 30° N) with a resolution of 0.1°. This is done using a conservative remapping via the remap function of the Climate Data Operators (CDO) version 1.9.5 (Schulzweida, 2019).

The size of the model output represents a challenge for the analysis. 30 days (corresponding to 240 timesteps) of one 3-hourly 3D field, interpolated to the 0.1° latitude-longitude grid covering only the tropics, have a size of about 150 Gigabytes. For 9 models and six variables this adds up to more than 8 TB. To reduce the amount of data we randomly sample a subset of grid points from each output timestep and only use the corresponding vertical profiles of each quantity for further analysis. Only grid points located over ocean are sampled. About 42,000 profiles are selected for each of the 240 timesteps, resulting in a total of 10 million profiles for each model. This roughly corresponds to 1% of the total number of tropical profiles over ocean. By repeating the random sampling several times for the same model we estimated the sampling uncertainty for the quantities analysed in this study to be negligibly small compared the inter-model differences we identify. Thus, the thinning of the data does not affect the results of this study.

The fifth generation of the ECMWF atmospheric reanalysis (ERA5) (Hersbach et al., 2020) serves as an observationally constrained reference data set in our comparison. It should be pointed out that potential biases with respect to observations exist in the ERA5 humidity fields. Xue et al. (2020) found a wet bias with respect to satellite observations in the free troposphere, which is most pronounced in regions of large-scale subsidence. Nevertheless, the dataset provides a valuable constraint of the humidity distribution and can be used to estimate its natural variability. Gridded atmospheric variables are provided at a spatial resolution of 31 km. We use 3-hourly output corresponding to the output times of the DYAMOND models and post-process it in the same way as the model output.

3 Humidity differences in DYAMOND models

In this section we quantify the differences in free-tropospheric humidity in the DYAMOND models, first in the tropical mean and subsequently in moisture space. Furthermore, we investigate whether the models' humidity anomalies are connected to anomalies in the resolved humidity transport.

3.1 Tropical mean

We focus on inter-model differences in relative humidity (RH) rather than absolute humidity, because they are a more direct measure of the radiative impact. The reason behind this is that differences in absolute humidity and temperature are positively correlated at constant RH, but their radiative impacts are counteractive and hence compensate to a large degree. This will be discussed in more detail in the second part of this paper.

RH is calculated for each of the randomly selected profiles and their associated values of q, p and T as RH $= \frac{e}{e_s(T)}$, where e is the water vapour pressure and $e_s(T)$ is its saturation value at temperature T. For $e_s(T)$ we take the value over water for T above the triple point T_t, the value over ice for T below $T_t - 23\,\mathrm{K}$. For intermediate T a a combination of both is used following the IFS documentation (ECMWF, 2018).

Overall, the models all capture the typical C-shape of the tropical mean RH profile with two maxima, one atop the boundary layer and one at the tropopause, and a minimum in the mid troposphere (Figure 1). The models' RH distributions also agree remarkably well with the ERA5 distribution. In fact, the multi-model mean RH (not shown) differs from ERA5 by less than 2% RH throughout the troposphere, except from the altitude region above 15 km.

Nevertheless, there are considerable differences among the models. The inter-model standard deviation σ(RH) (Figure 1c) has a distinct maximum around the top of the boundary layer (BL). The transition from the BL to the free troposphere is marked by a steep gradient in RH. Therefore, differences in the depth of the BL cause a large inter-model spread in RH. In IFS the humidity gradient at the top of the BL is particularly steep and the lower free troposphere is significantly dryer than in other models. Generally, in most models the BL is deeper than in ERA5. The inter-model spread is smallest in the mid troposphere between 4 and 10 km altitude. In that region σ(RH) is 2–3% RH and approximately constant with height. RH is lower than in ERA5 in the majority of models, except ICON and NICAM. Above 10 km σ(RH) increases with altitude and exceeds 8% RH at 100 hPa.

Anomalies in RH can either be caused by anomalies in absolute humidity (measured by q) or temperature T. In the DYAMOND models, T anomalies are small close to the surface, where they are constrained by identical SSTs, and increase with height throughout the free troposphere, where the temperature profile is set by convection and radiation (Figure 2a,b). In the lower and mid troposphere RH anomalies primarily reflect anomalies in q (Figure 1b, Figure 2d) and the impact of T anomalies on RH is small. In the upper troposphere, however, T anomalies gain influence. There, RH anomalies reflect both anomalies in T and q. Anomalies in T and q are highly correlated in the upper troposphere (Figure 2b,d), i.e. q is small in cold models and large in warm models. There, T differences are so large that differences in RH play a minor role in determining whether one model's absolute humidity is small or large as compared to another model's.

That the DYAMOND simulations were run just over one month (August/ September 2016) represents a potential limitation for the intercomparison, especially for variables that are subject to high internal variability on longer time scales. To estimate the internal variability of RH, we calculate the interannual variability in the mean August/September RH distribution based on five years (2014-2019) of the ERA5 reanalysis, shown as the dotted line in Figure 1c. Given that interannual variations in free-tropospheric humidity are primarily driven by SST variations (Chuang et al., 2010) and the five years include a strong El Niño event in 2015/2016, the interannual variability rather represents an upper bound for the internal variability one could expect in the DYAMOND runs with fixed SST. Despite this, the inter-model standard deviation is significantly larger than the ERA5 interannual variability throughout the troposphere, suggesting that the inter-model differences are systematic model biases rather than a result of poorly sampled internal variability.

To put the inter-model spread in DYAMOND into perspective, we compare it to the inter-model spread in 29 GCMs that participated in the AMIP experiment of the Coupled Model Intercomparison Project phase five (CMIP5) (Taylor et al., 2012). Like the DYAMOND simulations the AMIP simulations were run with prescribed SSTs. To make the comparison with the 30-day DYAMOND simulations as fair as possible only one August is selected from the AMIP simulations and tropical mean vertical profiles of RH are calculated for ocean regions only. Throughout most of the free troposphere σ(RH) in the DYAMOND ensemble is smaller by a factor of two and more compared to the CMIP5 AMIP ensemble (Figure 1c), indicating that the tropical mean free-tropospheric humidity distribution is better constrained in GSRMs. An exception is the lower free troposphere: the peak in σ(RH) at the top of the BL is less pronounced in CMIP5 than in DYAMOND, indicating that variations in the depth of the BL may be smaller in the CMIP5 models. However, part of the smaller spread in the CMIP5 models is also explained by the fact that the hydrolapse in these models is generally less steep, which is evident from the CMIP5 multi-model mean RH profile (Figure 1a). RH differences caused by a shift in the height of the hydrolapse are therefore smaller, but dispersed over a broader layer.

The reduced spread in free-tropopsheric RH in the DYAMOND ensemble is even more remarkable considering that the DYAMOND models were not tuned for this experiment. Many of them were even run in the storm-resolving configuration for the first time. However, as we will show in Section 4, the remaining humidity differences still have a significant impact on the clear-sky radiation budget.

3.2 Moisture space

To distinguish between different dynamic regimes of the tropics, which are not necessarily co-located in different models, we compare humidity statistics in moisture space (Bretherton et al., 2005; Schulz & Stevens, 2018; Naumann & Kiemle, 2020). To span the moisture space, the randomly selected atmospheric profiles (Section 2.2) are ranked by their vertically integrated water vapour (IWV). The integration is performed from the surface to an altitude of 20 km for all models.

Inter-model differences in the distribution of IWV are most pronounced at high IWV values (Figure 3). This is apparent when comparing different percentiles of IWV. While the 25th percentiles of all models lie within a range of 2.2 $\mathrm{kg\,m^{-2}}$, the 75th percentiles differ by more than 10 $\mathrm{kg\,m^{-2}}$ between the two most extreme models IFS and NICAM. The overall shape of the IWV distribution differs among models. For IFS and NICAM distributions are approximately uniform over a large range of IWV values, whereas the distribution of ARPEGE-NH has a pronounced peak at IWV values of about 50 $\mathrm{kg\,m^{-2}}$. For the remaining models (including ERA5) distributions are more bimodal with a first peak at 25–30 $\mathrm{kg\,m^{-2}}$ and a second peak at 50-55 $\mathrm{kg\,m^{-2}}$. The exact position and the

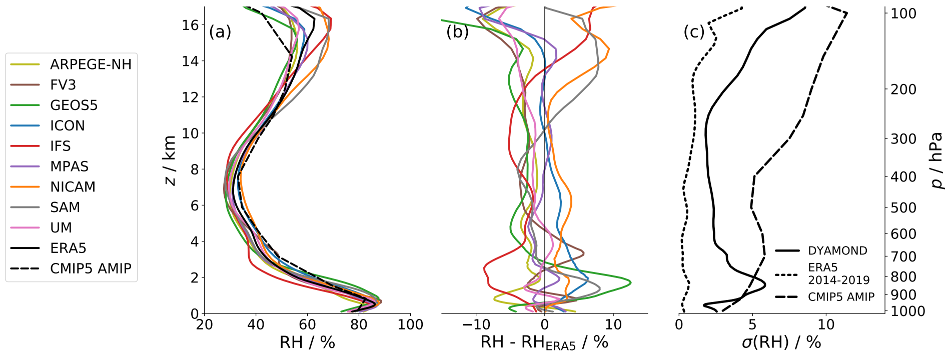

Figure 1. Tropical mean RH profiles and inter-model spread in the DYAMOND ensemble. (a) Tropical mean vertical profiles of RH over ocean regions from all DYAMOND models (colours), the ERA5 reanalysis (black solid) and the CMIP5 AMIP multi-model mean (black dashed). (b) Vertical RH profiles for the DYAMOND models shown as deviation from the ERA5 profile. (c) Inter-model standard deviation of tropical mean RH (solid line). For comparison the inter-annual spread in five years (2014-2019) of ERA5 (dotted line) as well as the inter-model spread in the CMIP5 AMIP ensemble (dashed line) are shown. For a representative comparison with DYAMOND only one August was selected from the CMIP5 AMIP runs.

relative strengths of the two peaks differ among the models. In SAM the first peak is particularly pronounced, whereas in ICON the second peak is comparably strong. Bimodality is a known feature of the IWV distribution over tropical oceans, which is not reliably reproduced by GCMs (Mapes et al., 2018). Our results indicate that this problem is similarly pronounced in GSRMs.

To display quantities in moisture space IWV-ranked profiles from each model are split into 50 blocks, each containing an equal amount of profiles corresponding to two percentiles of IWV. Quantities are then averaged over each block. Note that this block-averaging results in an x-axis that is linear in the percentile of IWV rather than in IWV itself. This also means that the comparison of different models in moisture space is made at a certain IWV percentile rather than a certain IWV value. In the multi-model mean the non-linear distribution of IWV values in moisture space is noticeable in the driest and moistest percentiles, respectively, where the increase in IWV is steeper than in the intermediate percentiles (Figure 4d). Again, it is apparent that the inter-model spread in IWV, which is indicated by the shading around the multi-model mean, increases from low to high percentiles.

SST increases from about 292 K in low IWV percentiles to about 302 K in high percentiles (Figure 4d). The SST gradient weakens from dry to moist regimes, similar to how the meridional SST gradient weakens from the subtropics towards the inner tropics. The inter-model standard deviation in block-averaged SSTs is around 0.15 K, implying that the the distribution of SST in moisture space is very similar among models. The underlying PDF of SSTs is identical in all models, which, compared to other quantities like IWV, puts an additional constraint on the SST distribution in moisture space.

Block-averaged vertical velocities (Figure 4c) indicate that the large-scale circulation is directed upward in the highest 5–10 IWV percentiles and downward in drier regions. The blocks with positive vertical velocities correspond to the regions of intense

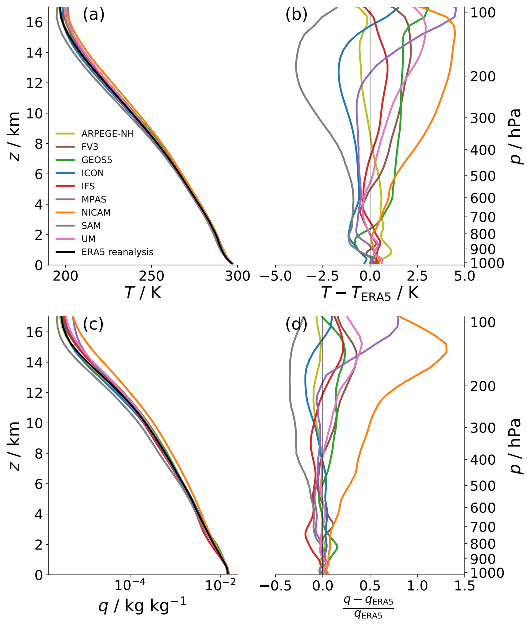

Figure 2. Tropical mean vertical profiles of specific humidity q and temperature T over ocean regions from all DYAMOND models. Vertical profiles of q (a, b) and T (c, d) shown as absolute values together with the ERA5 profiles (a, c) and as deviation from the ERA5 profiles (b, c). Deviations in q are in fractional units, i.e. normalized by the ERA5 value (q_{ERA5}).

rainfall in the Intertropical Convergenze Zone (ITCZ) in the deep tropics, where deep convection is concentrated. The drier blocks correspond to trade wind regimes. There, the free troposphere is characterized by large-scale subsidence, which increases in strength with decreasing IWV. At the transition from deep convective to subsidence regimes near the 90th IWV percentile vertical velocities are negative in the lower free troposphere and positive aloft. These blocks represent an advanced state in the life cycle of deep convection associated with upper-level anvil clouds. This state is characterized by ascent above the freezing level (which is located around 5 km) and descent below, driven by condensation and freezing above the freezing level, and melting and evaporation of precipitation below (Betts, 1990). The amount of high-level clouds increases from dry to moist regimes, as reflected by a sharp decrease in all-sky OLR in the moist blocks (Figure 4d).

The largest RH values are found in the BL (4a), where moisture is provided by evaporation from the surface. The RH in the BL is relatively constant throughout moisture space. Where air rises from the BL to the free troposphere in deep convective plumes it cools and its RH increases until saturation is reached. Therefore, the highest RH values in the free troposphere are found in deep convective regions. Saturated air detraining from deep convection moistens the surrounding regions corresponding to the lower-IWV blocks in moisture space. As detrainment from deep convection preferably takes place in the upper troposphere, a second maximum in RH is found there, losing in strength towards drier blocks. As the air subsides it warms and dries. The lowest RH values therefore occur in the free troposphere of the subsidence regions. Particularly in the high IWV percentiles a plateau in RH is apparent near the freezing level at around 5 km. Latent heat release from ice formation enhances the stability at this level, which causes deep convection to preferably detrain there (Stevens et al., 2017).

Displaying inter-model differences in moisture space reveals how they are distributed over the different regimes of the tropics. RH anomalies for individual models are shown in Figure A1 in Appendix A. Here we focus on the inter-model standard deviation σ(RH), shown in Figure 4b. First, it is apparent that the large inter-model spread in the upper troposphere (Figure 1) prevails throughout the entire tropics. In the tropopause region σ(RH) exceeds 10% RH everywhere except from the driest part of the subsidence regions. Second, the local maximum in σ(RH) at the top of the BL is most pronounced in the driest regimes, where the RH gradient between the BL and the free troposphere is steepest (Figure 4a). In moister regions, where the RH gradient is less steep, the maximum in σ(RH) is weaker but broader. Third, in the mid troposphere σ(RH) increases from less than 1% RH in the lowest IWV percentiles to more than 5% RH near the 90th percentile. The largest part of the spread in tropical mean mid-tropospheric RH stems from the region representing the transition from subsidence to deep convective regimes (cf. Figure 4c). In the moistest 5 percentiles of IWV the inter-model spread decreases again. In these regimes deep convection keeps the RH close to 100% in all models.

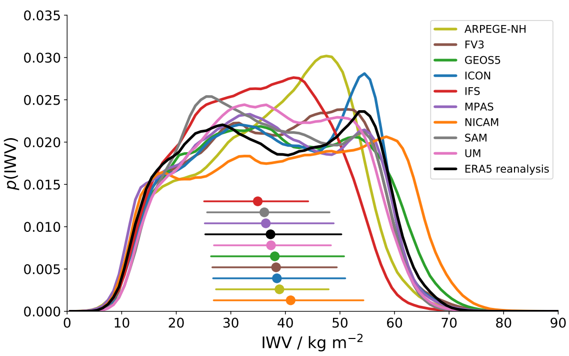

Figure 3. Probability density function of integrated water vapour (IWV) over tropical ocean regions in the DYAMOND models and ERA5. Percentiles of each model's are shown below the curves: Coloured circles indicate the medians of the distributions, horizontal bars range from the 25th to the 75th percentile.

3.3 Humidity transport by the resolved circulation

At this point an open question is which physical processes control the humidity differences in the DYAMOND ensemble. Besides the sub-grid-scale processes (i.e., phase change, turbulent mixing and radiation), that we cannot diagnose from the limited model output, transport by the resolved circulation has been suggested to play a major role (e.g Sherwood, 1996; Pierrehumbert & Roca, 1998; Dessler & Sherwood, 2000). As a step towards better understanding the physical causes behind the humidity differences, we investigate whether models with an anomalously high RH are associated with an anomalously high RH transport.

The tendency of RH due to resolved transport can be diagnosed from the model output:

$$\left(\frac{\partial \mathrm{RH}}{\partial t}\right)_{\mathrm{transport}} = \frac{\mathrm{RH}}{p}\frac{\mathrm{d}p}{\mathrm{d}t} - \mathrm{RH}\frac{1}{e_s}\frac{\mathrm{d}e_s}{\mathrm{d}T}\frac{\mathrm{d}T}{\mathrm{d}p}\frac{\mathrm{d}p}{\mathrm{d}t} - \vec{v}\cdot\nabla\mathrm{RH}, \tag{1}$$

where $\vec{v}$ denotes the three-dimensional velocity. The first two terms on the right hand side describe the change in RH caused by a pressure change following an air parcel. A

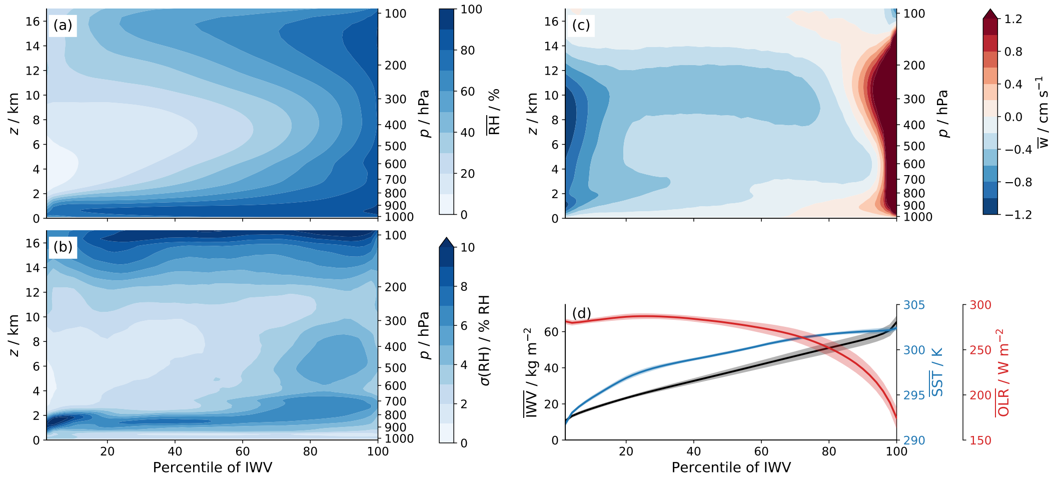

Figure 4. Distributions of different quantities in moisture space: (a) multi-model mean RH, (b) multi-model standard deviation of RH, (c) multi-model mean vertical velocity (d) multi-model mean IWV (black), SST (blue) and all-sky OLR (red). In (d) the inter-model standard deviation is denoted by shaded areas around the multi-model mean values.

change in pressure affects both the water vapour pressure e (first term) and the temperature T, which determines the saturation water vapour pressure e_s according to Clausius-Clapeyron (second term). We assume that $\frac{\mathrm{d}p}{\mathrm{d}t} = \vec{v} \cdot \nabla p$, so the pressure of an air parcel follows the environmental pressure, and that temperature changes adiabatically with pressure. Note that the second term generally dominates over the first one, so that a decrease in pressure in a rising air parcel causes its RH to increase. The third term on the right hand side denotes the advection of RH. For vertical transport it is the second term on the right hand side of Equation 1 that dominates the RH tendency, because pressure changes are large for a vertical motion. For horizontal transport, however, pressure changes are small and the third term (the advection term) is the dominant one.

We calculate the transport tendencies individually for each of the randomly selected profiles (Section 2.2) and then block-average them in moisture space. Horizontal gradients are calculated based on the selected profiles and their neighbouring profiles in physical space using central finite differences.

Figure 5 shows the RH transport tendencies for the multi-model mean in moisture space. The total tendency is shown in panel (c), the contributions from vertical and horizontal circulation are shown in (a) and (b), respectively. The vertical transport results in a strong moistening tendency in the highest IWV percentiles, which are characterized by positive vertical velocities (see also Figure 4), and in overall drying tendencies in the subsidence regions. Horizontal transport dries the moistest percentiles in the lower and mid troposphere. This drying is associated with the entrainment of dryer air from the surroundings in deep convective regimes. In the rest of the free troposphere the horizontal transport moistens the air, particularly in the upper troposphere, where the detrainment from deep convection takes place preferentially. From the total transport tendency it is clear that vertical and horizontal transport generally do not balance each other. Assuming that the RH distribution in moisture space is in a steady state, other (sub-grid-scale) processes must act to balance the transport. These include microphysical processes, turbulent mixing and radiation. We would expect RH tendencies due to microphysical processes to be most active in moist regimes, where condensation certainly acts to compensate a part of the moistening by the transport. An estimation of the RH tendency due to clear-sky radiative cooling indicates that this term is small (not shown) and plays a minor role in compensating the transport tendency.

To examine whether model anomalies in the transport tendencies are related to model anomalies in RH, we correlate them at each point in moisture space. A positive correlation indicates that models with high RH values are associated with an anomalously large transport tendency and vice versa at the respective point in moisture space. In that case, the transport anomalies would act to reinforce the RH anomalies. Where the correlation is negative, models with high RH are associated with a weak transport tendency. Our interpretation of this is that the unresolved processes act to reinforce the humidity anomaly, and the resolved transport, which balances those terms, has to compensate. A weak correlation indicates that the resolved transport is a process of minor importance and the actual balance is between other processes.

Positive correlations between anomalies in RH and anomalies in total transport mainly occur in the upper troposphere in the altitude region above 10 km (Figure 6). The total transport anomalies in this region are partly caused by anomalies in the vertical transport, which are associated with different representations of the Brewer-Dobson circulation in the models, and partly by anomalies in the horizontal transport (not shown). Positive correlations are also found at the edge of deep convective regimes in the altitude region between 7 and 10 km altitude, which is associated with anvil clouds (Section 3.2). Total transport anomalies there are mainly due to anomalies in vertical transport (not shown). Thus, in the anvil regions models with stronger vertical transport are moister. A small area of positive correlations also occurs in the lower free troposphere in the driest IWV percentiles. There, anomalies in the horizontal transport are the dominant ones. Throughout the rest of the free troposphere correlations are weak or negative, so transport anomalies do not act to reinforce the RH anomalies. A broad region of negative correlation is found in the mid troposphere in the anvil regions. A possible explanation could be that models with anomalously high RH in these regions are those with anomalously strong evaporation of precipitation (and vice versa). A stronger evaporative cooling causes stronger downdrafts and thereby also enhances the drying by vertical transport. Hence, the RH anomalies might be caused by differences in the microphysics, but the transport reacts to it, which can result in the negative correlation.

In summary, anomalies in the resolved transport can only explain RH anomalies in some regions, mostly in the upper troposphere above 10 km. Anomalies in the remaining parts of the free troposphere must be mainly related to other, unresolved processes. The most likely candidates are microphysical processes and turbulent/ shallow convective mixing.

4 Impact of RH anomalies on clear-sky OLR

To quantify the effect of the inter-model differences on the radiation balance, we translate them into differences in clear-sky OLR using a radiative transfer model. OLR differences are analysed in moisture space to determine how much different tropical moisture regimes contribute to the inter-model spread in tropical mean OLR. Furthermore, we investigate in which altitude regions humidity differences have the strongest impact on OLR. This allows us to identify the regions of the tropical troposphere in which a further reduction of humidity differences would be most beneficial.

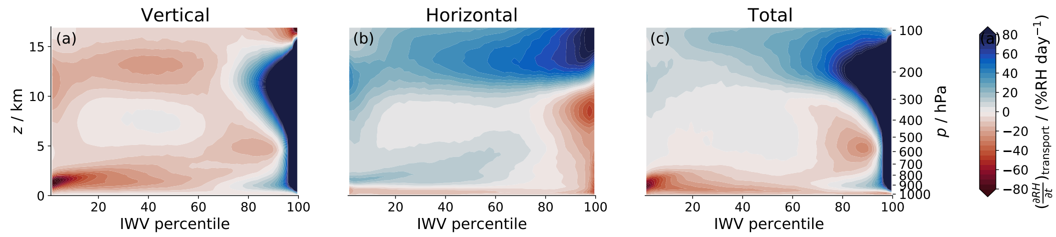

Figure 5. Multi-model mean RH tendencies due to (a) vertical, (b) horizontal and (c) total transport by the resolved circulation in moisture space.

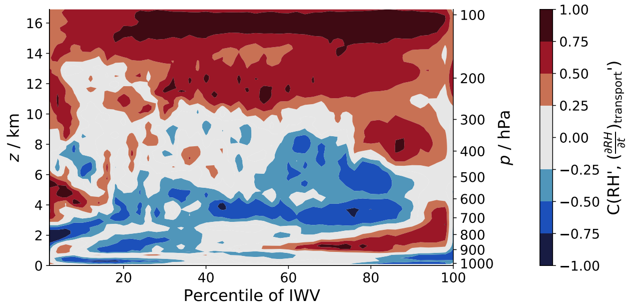

Figure 6. Correlation between model anomalies in RH and model anomalies in RH transport tendencies in moisture space. Positive correlations (red) indicate that models with high RH values are associated with an anomalously strong humidity transport and vice versa.

Fundamentally, clear-sky OLR is determined by the temperature of the surface and the temperature profile of the atmosphere as well as the concentration of greenhouse gasses in the atmosphere. The surface temperature plays a role in the window regions of the spectrum (between 800 to 1200 cm^{-1}), where the absorption by greenhouse gasses is weak and the radiation emitted from the surface can penetrate the atmosphere and directly escape to space. In the remaining parts of the spectrum the absorption of radiation by greenhouse gasses makes the atmosphere opaque, so that the radiation escaping at the TOA originates from the atmosphere rather than the surface.

For the OLR anomalies in the DYAMOND models we expect that anomalies in the surface temperature play a minor role, since SSTs are prescribed and their distributions in moisture space are very similar among models (Figure 4). Furthermore, in our OLR calculations we only consider the effect of model anomalies in water vapour. We expect the effect of differences in other greenhouse gasses to be small and therefore fix their concentrations in our radiative transfer simulations. Thus, in the DYAMOND models anomalies in clear-sky OLR are primarily caused by anomalies in atmospheric temperature and absolute humidity.

4.1 Radiative transfer simulations

The radiative transfer simulations to obtain clear-sky OLR are performed with the Rapid Radiative Transfer Model for GCMs (RRTMG Mlawer et al., 1997). RRTMG is is a well validated fast radiative transfer code used in various weather and climate models, also in several of the DYAMOND models. For this study we use RRTMG through the Python package konrad (DOI: 10.5281/zenodo.3899702), which in turn uses the CliMT

Python interface for RRTMG (Monteiro et al. 2018).

OLR is calculated based on the block-averaged profiles of pressure, temperature, and specific humidity in moisture space (Section 3.2). Calculating OLR from block-averaged profiles rather than from individual profiles induces an error, since radiation is non-linear in temperature and humidity. We found that OLR calculated from block-averaged profiles is generally lower than OLR calculated based on individual profiles. This is in line with the idea that fluctuations in humidity increase OLR (Pierrehumbert et al., 2007), so averaging out these fluctuations leads to a reduction of OLR. However, the resulting bias is very similar in all models, so that the effect on inter-model differences in OLR is negligible.

To characterize the surface we additionally select surface pressure and SST from the model output, in the same way as for the other variables (Section 2.2). The surface emissivity is assumed to be 1. For other gasses than water vapour we use fixed vertical profiles in accordance with those in Wing et al. (2017): The ozone volume mixing ratio follows a gamma distribution in pressure and vertically constant volume mixing ratios are assumed for O_2, CO_2, CH_4 and N_2O.

For the radiative transfer simulations we interpolate profiles from all models on a uniform vertical grid ranging from the surface to an altitude of 20 km with a resolution of 100 m. The top at 20 km corresponds to the maximum altitude for which output is available from all models. For our purpose OLR is defined as the longwave upward radiative flux at this level. Due to this definition the inter-model differences in OLR only reflect temperature and humidity differences in the troposphere, potential differences in the stratosphere are ignored. Note that due to the missing stratosphere the absolute value of the OLR defined at 20 km has a positive offset compared to the "true" OLR defined at a higher TOA. However, this is not relevant for our results since we are only interested in the effect of differences in the troposphere.

We focus only on the clear-sky case here, so any cloud condensate contained in the profiles is ignored. Clouds, particularly those at high altitudes, have a strong impact on OLR. Hence, model differences in cloud properties can cause significant differences in all-sky OLR, which are not considered here.

4.2 Model differences in clear-sky OLR

Tropical mean clear-sky OLR differs by more than 4 $\mathrm{Wm^{-2}}$ between the two most extreme models IFS and ICON (Figure 7a). The multi-model standard deviation in tropical mean clear-sky OLR is 1.2 $\mathrm{Wm^{-2}}$. This is a substantial spread given the fact that the climate forcing due to a doubling of CO_2 is about 3.7 $\mathrm{Wm^{-2}}$ (Collins et al., 2013). In some models, e.g. UM and ARPEGE-NH, both positive and negative OLR anomalies occur across moisture space, which partly cancel in the tropical mean.

Two moisture regimes stand out due to a particularly large spread in clear-sky OLR (Figure 7b): One local maximum in σ(OLR) occurs in rather moist regimes around the 80th percentile of IWV. This corresponds to the region at the transition from deep convective to subsidence regimes, where the inter-model RH spread in the mid troposphere maximizes (Figure 4b). A second, slightly weaker maximum in σ(OLR) is located at the dry end of moisture space. In the next section we aim to better understand why the OLR spread maximizes in these two regimes and which altitude regions in the troposphere contribute most.

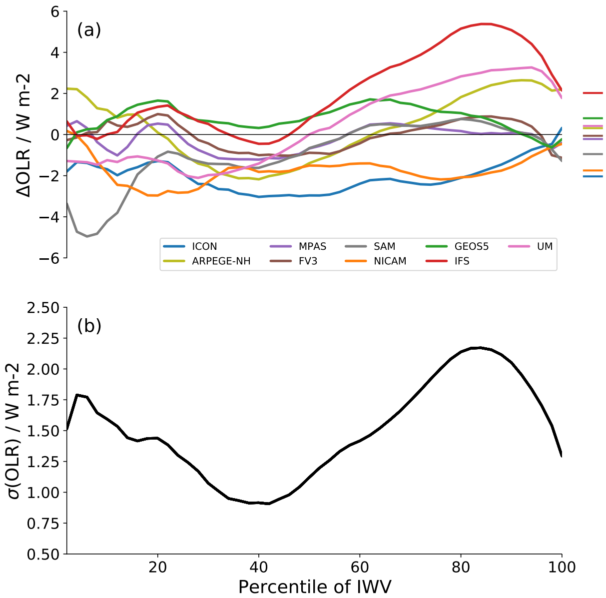

Figure 7. Inter-model differences in clear-sky OLR in moisture space. (a) Anomalies in clear-sky OLR for each model, defined as the deviation from the ERA5 value and (b) inter-model standard deviation of clear-sky OLR.

4.3 Radiative kernels

To examine how different regions in moisture space contribute to the spread in tropical mean clear-sky OLR, for each of the 50 blocks in moisture space we decompose each model's OLR anomaly into contributions from individual atmospheric layers using the radiative kernel method (Soden et al., 2008).

Dividing the atmosphere into N vertical layers and linearising around the ERA5 state that we use as reference state, a model's OLR anomaly ΔOLR can be written as:

$$\Delta\mathrm{OLR} \approx \sum_{i=1}^{N} \left(K_i^e \, \Delta \mathrm{e}_i + K_i^T \, \Delta \mathrm{T}_i \right) \approx \sum_{i=1}^{N} K_i^{\mathrm{RH}} \, \Delta \mathrm{RH}_i. \tag{2}$$

Here, the index i denotes the vertical layer. The vectors $\mathbf{K^x}$ are radiative kernels that describe the sensitivity of OLR to changes in a variable x in each layer:

$$K_i^x = \frac{\partial \mathrm{OLR}}{\partial \mathrm{x}_i}. \tag{3}$$

The first approximation in Equation 2 assumes that OLR anomalies are primarily caused by anomalies in atmospheric T and e, the effect of anomalies in surface temperature is assumed to be negligible. Moreover, it is assumed that contributions from each layer to the OLR response are independent, neglecting potential masking effects. For example, when a model has a strong positive e anomaly in the upper troposphere, this would increase the optical thickness of the atmosphere there and thereby weaken the effect of anomalies below. Despite these assumptions the kernels $\mathbf{K^e}$ and $\mathbf{K^T}$ can be used to approximate the OLR anomalies of the DYAMOND models with good accuracy, which is shown in Figure B1 in Appendix B. In Appendix B we also describe the calculation of the kernels.

Perturbations in e and T have opposite effects on OLR, which is evident from the different signs of the respective kernels (Figure B1). Increasing temperature in a given atmospheric layer increases the emission from that layer and hence increases OLR. Conversely, increasing absolute humidity in a given layer reduces OLR both by shifting the effective emission layer upwards to colder temperatures in the water vapour bands, and by closing the atmospheric window. At constant RH perturbations in e and T are positively correlated, so their effects on OLR compensate to some degree. It is well known that in the water vapour bands, the spectral regions in which the water vapour optical depth is larger than 1, modulo foreign broadening the emission from a layer to space depends only on RH (Nakajima et al., 1992; Ingram, 2010). This behaviour is often referred to as "Simpsonian", as it has been recognized since the early work of Simpson (1928). Therefore, we can assume that OLR anomalies in the DYAMOND models are primarily determined by RH anomalies. This corresponds to the second approximation in Equation 2.

A perturbation in RH can be produced isothermally, i.e. by varying e and keeping T constant, or isobarically, i.e. by varying T and keeping e constant. Therefore, there are two ways to define a RH kernel, which we refer to as $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ and $\mathbf{K}^{\mathrm{RH},\mathbf{T}}$, respectively:

$$\begin{aligned} K_i^{\mathrm{RH},e} &= \frac{\partial \mathrm{OLR}}{\partial \mathrm{RH}_i}\bigg|_{T=\mathrm{const.}} &&= e_s\, K_i^e \\ K_i^{\mathrm{RH},T} &= \frac{\partial \mathrm{OLR}}{\partial \mathrm{RH}_i}\bigg|_{e=\mathrm{const.}} &&= -\frac{e_s}{\mathrm{RH}}\left(\frac{de_s}{dT}\right)^{-1} K_i^T. \end{aligned} \tag{4}$$

To translate $\mathbf{K^e}$ and $\mathbf{K^T}$ into RH kernels they have to be weighted by a factor describing the change of RH for a change in e or T, respectively. For $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ this factor is equal to the saturation water vapour pressure e_s. For $\mathbf{K}^{\mathrm{RH},\mathbf{T}}$ the dependence of e_s on T given by the Clausius Clapeyron relation has to be taken into account. $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ and $\mathbf{K}^{\mathrm{RH},\mathbf{T}}$ are identical to the extent that the OLR response to a given change in RH is independent of whether this change is produced by a change in e or in T.

Differences between $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ (Figure 8a) and $\mathbf{K}^{\mathrm{RH},\mathbf{T}}$ (Figure B2) indicate that to a certain degree it does matter whether a RH perturbation is caused by a perturbation in e or in T. We elaborate a bit more on these differences in Appendix B. As evident from comparing Figure 8c and Figure B2c, OLR anomalies approximated using $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ are more accurate than those approximated using $\mathbf{K}^{\mathrm{RH},\mathbf{T}}$. This implies that RH anomalies in the DYAMOND models are primarily caused by anomalies in absolute humidity rather than temperature (at least in the altitude regions that are most relevant for OLR). Therefore, for the further analysis we concentrate on $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$.

Overall, there is good agreement between true (directly calculated) OLR anomalies and those approximated with Equation 2 using $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ (Figure 8c). The largest deviations occur for ICON, for SAM in the lowest IWV percentiles and for ARPEGE-NH in moist percentiles. The inter-model standard deviation σ(OLR) is well reproduced with the approximated OLR (Figure 8d), except from the lowest IWV percentiles, where it is slightly underestimated. This is mainly caused by the deviations in SAM and ICON. For most models the approximation from RH anomalies is slightly less accurate than the one from e and T anomalies (cf. Figure B1). An exception is NICAM, for which the OLR approximated from RH anomalies matches the true OLR much better than the one approximated from e and T anomalies. Overall, we conclude that inter-model differences in RH indeed explain a major part of the differences in clear-sky OLR in the DYAMOND models.

4.4 Relative importance of different altitude regions

The impact of RH anomalies for the radiation budget is determined by the magnitude of the anomalies and the sensitivity of OLR to a given perturbation in RH, which is described by the radiative kernel $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ (Equation 2). $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ is negative throughout the tropical troposphere (Figure 8a), indicating that an increase in RH leads to a decrease in OLR. Its absolute value is largest in the mid troposphere in the dry subsidence regimes. The reason for this can be understood from Equation 4, which states that $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ is equal to the product of $\mathbf{K}^{\mathbf{e}}$ and e_s. $\mathbf{K}^{\mathbf{e}}$ generally increases with height and from moist to dry regimes (Figure B1). This is due to changes in the degree of saturation in the water vapour bands. In regions with low absolute humidity, i.e. in the upper troposphere and in dry regimes, absorption bands are radiatively less saturated, so the sensitivity to humidity changes is larger than for regions with high absolute humidity. At the same time e_s decreases with altitude. Hence, the product of $\mathbf{K}^{\mathbf{e}}$ and e_s maximizes in the mid troposphere of the dry regimes.

In low IWV percentiles there is a pronounced peak in $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ at an altitude of around 6 km. The peak weakens from dry to moist regimes as the absorption bands become more saturated. A very similar behaviour was found by Spencer and Braswell (1997) for base states with RH values roughly corresponding to those in the dry half of moisture space. For the moist half of moisture space we find that lower atmospheric layers (below 5 km) become relatively more important. A possible explanation for this could be the continuum absorption in the major atmospheric window region (approximately 800 to 1200 cm^{-1}), which acts to decrease the surface component of OLR as humidity increases in the lower troposphere. In contrast to the water vapour bands, saturation effects do not play a role for the continuum absorption (Allan et al., 1999). As a consequence, humidity perturbations in the lower troposphere become relatively more important for base states with high RH.

The product of the RH response kernel $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ and the RH inter-model standard deviation σ(RH) (Figure 8b) indicates where the actual inter-model differences have the strongest effect on clear-sky OLR. First, the top of the BL stands out as a narrow region of strong impact. OLR is not particularly sensitive to RH perturbations there (Figure 8a), but the inter-model differences in RH are large (Figure 4b) because the models differ in the depth of the BL. RH differences in a broad layer in the mid troposphere also significantly affect OLR. Integrated over its full width, the contribution from this layer is larger than that from the BL top. The mid troposphere is characterized by an increasing RH spread from dry to moist regimes with a pronounced maximum near the 80th IWV percentile (Figure 4b) and a decreasing OLR sensitivity from dry to moist regimes (Figure 8a). The combination of both results in a relatively uniform importance of RH differences across moisture space, with two local maxima occurring near the 30th and near the 80th IWV percentile. The layer over which RH differences have a considerable impact on OLR generally extends to higher altitudes in the dry regimes than in the moist regimes, which is a consequence of the more saturated water vapour bands in the moist regimes. Due to the low OLR sensitivity in the upper troposphere (above about 10–12 km) the large inter-model RH differences there (Figure 4b) have virtually no effect on OLR.

Not considering clouds has an effect on the response kernels. Particularly high clouds are important, because they mask some of the effect of temperature and humidity in lower atmospheric levels (Soden et al., 2008). They are mainly present in moist regimes, starting around the 60th IWV percentile in most models (not shown). Therefore, in these regimes we would expect the OLR sensitivity to RH perturbations to become stronger in the levels in which clouds are most abundant (roughly 8-12 km height) and weaker at lower levels. This could dampen some of the effect of the large RH differences in the lower and mid free troposphere in the moist regimes.

An important point to note is that the vertical integration of the product of $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ and $\sigma(\mathrm{RH})$, shown as the grey line in Figure 8d, does not yield the inter-model standard deviation in OLR, but a higher value, which is more uniform throughout moisture space. In many models RH anomalies have different signs in different altitude regions (Figure 1 and Figure A1). This information is not contained in $\sigma(\mathrm{RH})$. The effects of such opposite RH anomalies on OLR compensate to some degree. Interestingly, such compensating errors play a bigger role in the dry regimes, as indicated by the larger difference between the grey and the black line in Figure 8d and evident from Figure A1. In fact, it is only due to these effects that dry regimes contribute less to tropical mean differences in clear-sky OLR than moist regimes.

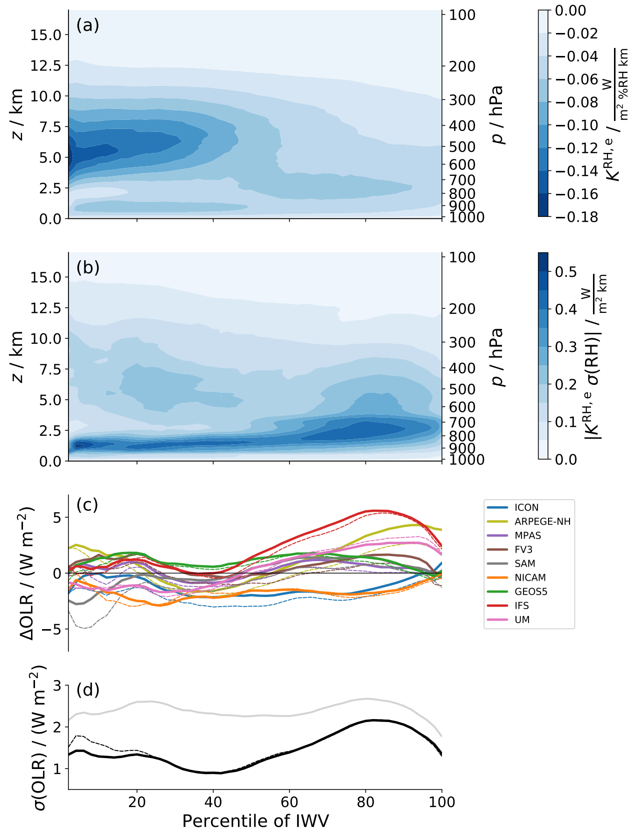

Figure 8. Impact of RH differences on clear-sky OLR in moisture space. (a) RH response kernel $\mathbf{K}^{\mathrm{RH,e}}$ showing the sensitivity of clear-sky OLR to a 1% change in RH in a 1 km layer under constant temperature for 50 blocks in moisture space, (b) inter-model standard deviation $\sigma(\mathrm{RH})$ weighted with $\mathbf{K}^{\mathrm{RH,e}}$, (c) OLR anomalies approximated from $\mathbf{K}^{\mathrm{RH,e}}$ and the RH anomalies of each model and (d) inter-model standard deviation in the approximated OLR. Thin dashed lines in (c) and (d) correspond to OLR calculated directly from temperature and humidity profiles (same as in Figure 7). The vertical integral of (b) is shown as the grey line in (c).

5 Summary and conclusions

In this study we quantified inter-model differences in tropical free-tropospheric humidity in an ensemble of nine different GSRMs that took part in DYAMOND, the first intercomparison project for models of this type. We focused on the effect of the humidity differences on the radiation budget and therefore concentrated on differences in RH rather than absolute humidity. The RH is most informative because in a large part of the spectrum the emission from a layer to space depends primarily on RH (Nakajima et al., 1992; Ingram, 2010).

We find that the inter-model spread in tropical RH in DYAMOND is reduced by about a factor of two compared to the CMIP5 AMIP ensemble, confirming that the RH distribution and hence the clear-sky OLR are better constrained at storm-resolving resolutions. A question that cannot be answered from the relatively short DYAMOND simulations is whether the spread in the water vapour feedback is also reduced in GSRMs. However, there are some reasons to be optimistic about this. On the one hand, to the extent that the water vapour feedback depends on the base-state RH, reducing the inter-model spread in present-day RH should also reduce the spread in the feedback. Evidence for the existence of such a state-dependence was recently found by Bourdin et al. (2021). On the other hand, the water vapour feedback depends on how much RH changes under warming. For the CMIP5 models it is differences in the RH response that contribute most to the spread in the water vapour feedback (Vial et al., 2013). However, given that the present-day RH is better constrained in GSRMs, it seems unlikely that the spread in the RH response increases.

Although humidity differences are reduced in the DYAMOND ensemble, they still cause a considerable spread of $1.2\,\mathrm{Wm^{-2}}$ in tropical mean clear-sky OLR. To better understand how different tropical moisture regimes contribute to this spread, it has proven useful to compare model fields in moisture space, i.e. sorted from low to high IWV. Combining the inter-model standard deviation $\sigma(\mathrm{RH})$ with radiative kernels (the sensitivity of OLR to RH perturbations) in moisture space allowed us to examine the radiative impact of the RH differences in a given dynamic regime and altitude region and hence to assess in which regions a further reduction would be most beneficial. Based on the results we can split the tropical free troposphere into four main regions:

1. The transition between the BL and the free troposphere. Throughout the tropics this altitude region (around 2 to 3 km) is characterized by a local maximum in the inter-model RH spread, with $\sigma(\mathrm{RH})$ exceeding 6% RH. These differences are associated with differences in the depth of the BL. Due to their large magnitude they contribute considerably to the spread in clear-sky OLR, although the sensitivity of OLR to a given RH perturbation is rather small in this altitude region.
2. The mid troposphere of moist regimes. This region ranges from about 3 km to 10 km in altitude and roughly covers the highest 50 percentiles of IWV in moisture space. With $\sigma(\mathrm{RH})$ up to 6% RH the inter-model spread in these moist regimes is substantially larger than in the same altitude region of dry regimes. The spread maximizes at the transition from deep convective to subsidence regimes near the 90th percentile of IWV. Although the OLR sensitivity to RH perturbations is moderate, the large RH differences cause the inter-model OLR spread to maximize in this region.
3. The mid troposphere of dry regimes. In this region the model agreement in RH is remarkably good. The inter-model standard deviation $\sigma(\mathrm{RH})$ is 1–3% RH and hence less than half of the standard deviation in moist regimes. However, the sensitivity of OLR to RH perturbations is considerably larger. Therefore, the small RH differences in the dry regimes have a comparable effect on clear-sky OLR as the larger differences in the moist regimes. This is why the inter-model OLR spread has a second, albeit slightly weaker local maximum in the dry regimes. The maximum is weaker than the one in the moist regimes because compensating effects due to opposite RH anomalies at different altitude regions occur more frequently in the dry regimes. The reason for this is not obvious and needs further investigation.
4. The upper troposphere. In the altitude region above 10 km the inter-model spread is generally large, with $\sigma(\mathrm{RH})$ exceeding 8% near the tropopause. However, the OLR sensitivity to RH perturbations is so small that the impact of these differences on the clear-sky OLR is negligible.

Our results are limited to the clear-sky case. High clouds, which are most abundant in the moist regimes, potentially mask some of the clear-sky effect (e.g. Soden et al., 2008) and hence reduce the radiative impact of the humidity differences in the mid troposphere. This highlights even more the importance of the dry regimes, where high clouds are rare.

As a step towards better understanding the physical causes behind the RH differences, we investigated whether RH anomalies are related to anomalies in transport by the resolved circulation. We find that anomalies in the resolved transport can only explain parts of the RH anomalies in the upper troposphere but not in the regions below, which are more relevant for the radiation budget. This suggests that sub-scale processes like microphysics and turbulence as well as their interaction with the large-scale circulation play a major role in controlling the differences in the most critical regions. This result does not contradict earlier studies, which emphasize the important role of the large-scale transport in setting the humidity distribution of the free troposphere (e.g. Sherwood, 1996; Pierrehumbert & Roca, 1998; Dessler & Sherwood, 2000). After all, the DYAMOND models all reproduce the basic shape of the RH distribution. Nevertheless, differences in the representation of sub-scale processes can cause subtle modifications in the RH distribution that manifest as inter-model differences.

We conclude that to further constrain the radiation budget in GSRMs it is most crucial to reduce the RH differences at the top of the BL and in the mid troposphere. Reducing the former by adjusting the depth of the BL seems possible with the current level of knowledge. Also, one would expect clear benefits from increased vertical resolution when it comes to representing the BL depth. On the other hand, observational reference data are sparse because satellite capacities to probe the BL region are still limited. Reducing the differences in the mid troposphere seems more challenging and requires a detailed understanding of how sub-scale processes affect the RH in these regions remote from deeper convection. An advantage is that this altitude region of the tropical atmosphere is extensively observed by satellites.

Appendix A RH anomalies in individual models

In Section 3.2 we focused on the inter-model spread in RH expressed by the inter-model standard deviation $\sigma(\mathrm{RH})$. Here we show how the RH deviates from ERA5 in moisture space for individual models (Figure A1). It is evident that for many models, particularly for ICON, NICAM and IFS, the largest part of the RH anomalies in the mid troposphere that are apparent in the tropical mean (Figure 1) stems from rather moist regimes. Furthermore, in all models RH anomalies of opposite sign exist at different altitude regions and across moisture space. As mentioned in Sections 4.2 and 4.4 their effects on tropical mean clear-sky OLR partly compensate. For example, the GEOS5 model has both an anomalously moist lower free troposphere (due to an anomalously deep BL) and an anomalously dry mid free troposphere in regions of intermediate IWV (Figure A1d). Due to the compensation of these opposite effects the OLR anomaly in these regions is rather small (Figure 7). In the UM model the lower and mid free troposphere are anomalously moist in dry regimes and anomalously dry in moist regimes (Figure A1j). The resulting OLR anomalies almost fully compensate in the tropical mean (Figure 7).

Appendix B Radiative kernels for water vapour pressure, temperature and relative humidity

To obtain the radiative kernels $\mathbf{K^e}$ and $\mathbf{K^T}$ for a given block in moisture space, OLR is calculated for the averaged ERA5 profiles in this block using the setup described in Section 4.1. The calculation is repeated with a small perturbation applied to e or T in one atmospheric layer, yielding the element of $\mathbf{K^e}$ of $\mathbf{K^T}$, respectively, for that layer. This is done successively for all layers. We perturb e by 5% of its absolute value and T by 1 K. The results are not sensitive to the exact size of the perturbation.

The kernels $\mathbf{K^e}$ and $\mathbf{K^T}$ can be used together with anomalies in e and T to approximate anomalies in clear-sky OLR (Equation 2) in the DYAMOND models with good accuracy (Figure B1e). The approximation is least accurate for the NICAM model. NICAM is the model with the largest anomalies in absolute humidity (Figure 2), so the assumption of linearity around the reference state starts to lose validity. In other models some smaller inaccuracies occur particularly in the dry half of moisture space. Most of them can be explained by SST anomalies that are not considered in Equation 2. Such SST anomalies have a stronger impact in the dry regions because the surface component of OLR is larger there than in moist regions. The largest deviations between true and approximated OLR anomalies in dry regimes arise for SAM and ARPEGE-NH. These are only partly explained by SST anomalies, so non-linearity or masking effects might play a role.

As explained in Section 4.3, OLR anomalies can also be approximated from RH anomalies and a RH kernel (Equation 2). There are two ways to define a RH kernel by varying either e or T (Equation 4), which we refer to as $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ and $\mathbf{K}^{\mathrm{RH},\mathbf{T}}$, respectively. Our main analysis above is based on $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ because it approximates the OLR anomalies more accurately. For completeness Figure B2 shows $\mathbf{K}^{\mathrm{RH},\mathbf{T}}$ and the OLR anomalies approximated using this version of the RH kernel. Compared to $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ (Figure 8a), $\mathbf{K}^{\mathrm{RH},\mathbf{T}}$ (Figure B2a) takes on larger absolute values (note the different colour scales in Figures 8 and B2), i.e. a 1% increase in RH causes a larger decrease in OLR if it is produced by decreasing T rather than increasing e. Furthermore, the peak altitude in $\mathbf{K}^{\mathrm{RH},\mathbf{T}}$ is lower than in $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$. These differences indicate that for OLR it does matter to a certain degree whether a RH perturbation is caused by a perturbation in e or in T. Nevertheless, considering that the physical mechanisms behind a change in OLR are very different for changes in e and T, the two kernels agree remarkably well, again demonstrating that the atmosphere behaves partly ”Simpsonian” (see Section 4.3).

Acknowledgments

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany‘s Excellence Strategy – EXC 2037 'CLICCS - Climate, Climatic Change, and Society' – Project Number: 390683824, contribution to the Center for Earth System Research and Sustainability (CEN) of Universität Hamburg.

DYAMOND data management was provided by the German Climate Computing Center (DKRZ) and supported through the projects ESiWACE and ESiWACE2. The projects ESiWACE and ESiWACE2 have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements No 675191 and 823988. The authors would like to thank the European Centre for Medium-Range Weather Forecasts (ECMWF) for providing the ERA5 data, which is available at the Copernicus Climate Change Service Climate Data Store (CDS) `https://cds.climate.copernicus.eu/cdsapp#!/home`. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups for producing and making available their model output. The CMIP5 AMIP data were accessed through DKRZ.

Version v0.8.0 of konrad is available at `https://github.com/atmtools/konrad/tree/v0.8.0`

We would like to thank Daniel Klocke for technical help and Lukas Kluft for technical help and valuable comments on the draft.

The authors declare not conflict of interest.

References

Allan, R. P., Shine, K. P., Slingo, A., & Pamment, J. A. (1999). The dependence of clear-sky outgoing long-wave radiation on surface temperature and relative humidity. *Quarterly Journal of the Royal Meteorological Society*, *125*(558), 2103–2126. doi: 10.1002/qj.49712555809

Betts, A. (1990). Greenhouse warming and the tropical water budget. *Bulletin of the American Meteorological Society*, *71*, 1464–1465.

Bourdin, S., Kluft, L., & Stevens, B. (2021). Dependence of climate sensitivity on the given distribution of relative humidity. *Earth and Space Science Open Archive*, 13. Retrieved from `https://doi.org/10.1002/essoar.10506044.1` doi: 10.1002/essoar.10506044.1

Bretherton, C., Blossey, P., & Khairoutdinov, M. (2005). An energy-balance analysis of deep convective self-aggregation above uniform SST. *Journal of the Atmospheric Sciences*, *62*(12), 4273–4292. doi: 10.1175/JAS3614.1

Brogniez, H., Roca, R., & Picon, L. (2005). Evaluation of the distribution of subtropical free tropospheric humidity in AMIP-2 simulations using METEOSAT water vapor channel data. *Geophysical Research Letters*, *32*(19), n/a–n/a. doi: 10.1029/2005GL024341

Bryan, G. H., Wyngaard, J. C., & Fritsch, J. M. (2003). Resolution requirements for the simulation of deep moist convection. *Monthly Weather Review*, *131*(10), 2394–2416. doi: 10.1175/1520-0493(2003)131⟨2394:RRFTSO⟩2.0.CO;2

Chuang, H., Huang, X., & Minschwaner, K. (2010). Interannual variations of tropical upper tropospheric humidity and tropical rainy-region SST: Comparisons between models, reanalyses, and observations. *Journal of Geophysical Research*, *115*(D21). doi: 10.1029/2010JD014205

Collins, M., Knutti, R., Arblaster, J., Dufresne, J., Fichefet, T., Friedlingstein, P., ... others (2013). Climate change 2013: the physical science basis.

Dessler, A., & Sherwood, S. (2000). Simulations of tropical upper tropospheric humidity. *Journal of Geophysical Research*, *105*(D15), 20155–20163. doi: 10.1029/2000JD900231

ECMWF. (2018). Ifs documentation cy45r1. In (chap. Part IV : Physical processes). Retrieved from `https://www.ecmwf.int/node/18714`

Harries, J. (1997). Atmospheric radiation and atmospheric humidity. *Quaterly Journal of the Royal Meteorological Society*, *123*(544), 2173–2186. doi: 10.1002/qj.49712354402

Held, I., & Soden, B. (2000). Water vapour feedback and global warming. *Annual Review of Energy and the Environment*, *25*(1), 441–475. doi: 10.1146/annurev.energy.25.1.441

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. doi: 10.1002/qj.3803

Ingram, W. (2010). A very simple model for the water vapour feedback on climate change. *Quarterly Journal of the Royal Meteorological Society*, *136*(646), 30–40. doi: 10.1002/qj.546

Jiang, J. H., Su, H., Zhai, C., Perun, V. S., Genio, A. D., Nazarenko, L. S., ... Stephens, G. L. (2012). Evaluation of cloud and water vapor simulations in CMIP5 climate models using NASA "a-train" satellite observations. *Journal of Geophysical Research: Atmospheres*, *117*(D14), D14105. doi: 10.1029/2011JD017237

John, V. O., & Soden, B. J. (2007). Temperature and humidity biases in global climate models and their impact on climate feedbacks. *Geophysical Research Letters*, *34*(18), L18704. doi: 10.1029/2007GL030429

Mapes, B. E., Chung, E. S., Hannah, W. M., Masunaga, H., Wimmers, A. J., & Velden, C. S. (2018). The meandering margin of the meteorological moist tropics. *Geophysical Research Letters*, *45*(2), 1177–1184. doi: 10.1002/2017GL076440

Miyamoto, Y., Kajikawa, Y., Yoshida, R., Yamaura, T., Yashiro, H., & Tomita, H. (2013). Deep moist atmospheric convection in a subkilometer global simulation. *Geophysical Research Letters*, *40*(18), 4922–4926. doi: 10.1002/grl.50944

Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., & Clough, S. A. (1997). Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research: Atmospheres*, *102*(D14), 16663–16682. doi: 10.1029/97JD00237

Nakajima, S., Hayashi, Y.-Y., & Abe, Y. (1992). A study on the "runaway greenhouse effect" with a one-dimensional radiative–convective equilibrium model. *Journal of the Atmospheric Sciences*, *49*(23), 2256–2266. doi: 10.1175/1520-0469(1992)049⟨2256:ASOTGE⟩2.0.CO;2

Naumann, A., & Kiemle, C. (2020). The vertical structure and spatial variability of lower-tropospheric water vapor and clouds in the trades. *Atmospheric Chemistry and Physics*, *20*(10), 6129–6145. doi: 10.5194/acp-20-6129-2020

Pierce, D., Barnett, T., Fetzer, E., & P.J., G. (2006). Three-dimensional tropospheric water vapor in coupled climatemodels compared with observations from the airs satellite system. *Geophysical Research Letters*, *33*(21), L21701. doi: 10.1029/2006GL027060

Pierrehumbert, R., Brogniez, H., & Roca, R. (2007). The general circulation. In T. Schneider & A. Sobel (Eds.), (pp. 143–185). Princeton University Press, Princeton, NJ.

Pierrehumbert, R., & Roca, R. (1998). Evidence for control of atlantic subtropical humidity by large scale advection. *Geophysical Research Letters*, *25*(24), 4537–4540. doi: 10.1029/1998GL900203

Satoh, M., Stevens, B., Judt, F., Khairoutdinov, M., Lin, S.-J., Putman, W., & Düben, P. (2019). Global cloud-resolving models. *Current Climate Change Reports*, *5*(3), 172–184. doi: 10.1007/s40641-019-00131-0

Schulz, H., & Stevens, B. (2018). Observing the tropical atmosphere in moisture space. *Journal of the Atmospheric Sciences*, *75*(10), 3313–3330. doi: 10.1175/JAS-D-17-0375.1

Schulzweida, U. (2019). *Cdo user guide.* Zenodo. doi: 10.5281/zenodo.3539275

Sherwood, S. (1996). Maintainance of the free-tropospheric tropical water vapor distribution. part ii: Simulation by large-scale advection. *Journal of Climate*, *9*(11), 2919–2934. doi: 10.1175/1520-0442(1996)009⟨2919:MOTFTT⟩2.0.CO;2

Sherwood, S., Roca, R., Meckwerth, T., & Andronova, N. (2010). Tropospheric water vapor, convection and climate. *Reviews of Geophysics*, *48*, RG2001. doi: 10.1029/2009RG000301

Simpson, G. (1928). Some studies in terrestrial radiation. *Memoirs of the Royal Meteorological Society*, *2*(16), 69–95.

Soden, B., Held, I., Colman, R., Shell, K., Kiehl, J., & Shields, C. (2008). Quantifying climate feedbacks using radiative kernels. *Journal of Climate*, *21*(14), 3504–3520. doi: 10.1175/2007JCLI2110.1

Soden, B., Jackson, D., Ramaswamy, V., Schwarzkopf, M. D., & Huang, X. (2005). The radiative signature of upper tropospheric moistening. *Science*, *310*(5749), 841–844. doi: 10.1126/science.1115602

Spencer, R., & Braswell, W. (1997). How dry is the tropical free troposphere? implications for global warming theory. *Bulletin of the American Meteorological Society*, *78*(6), 1097–1106. doi: 10.1175/1520-0477(1997)078⟨1097:hdittf⟩2.0.co;2

Stevens, B., Acquistapace, C., Hansen, A., Heinze, R., Klinger, C., Klocke, D., ... Zängl, G. (2020). The added value of large-eddy and storm-resolving models for simulating clouds and precipitation. *Journal of the Meteorological Society of Japan. Ser. II*, *98*(2), 395–435. doi: 10.2151/jmsj.2020-021

Stevens, B., & Bony, S. (2013). What are climate models missing? *Science*, *340*(6136), 1053–1054. doi: 10.1126/science.1237554

Stevens, B., Brogniez, H., Kiemle, C., Lacour, J.-L., Crevoisier, C., & Kiliani, J. (2017). Structure and dynamical influence of water vapor in the lower tropical troposphere. *Surveys in Geophysics*, *38*(6), 1371–1397. doi: 10.1007/s10712-017-9420-8

Stevens, B., Satoh, M., Auger, L., Biercamp, J., Bretherton, C. S., Chen, X., … Zhou, L. (2019). DYAMOND: the DYnamics of the atmospheric general circulation modeled on non-hydrostatic domains. *Progress in Earth and Planetary Science*, *6*(1). doi: 10.1186/s40645-019-0304-z

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*(4), 485–498. doi: 10.1175/BAMS-D-11-00094.1

Vial, J., Dufresne, J.-L., & Bony, S. (2013). On the interpretation of inter-model spread in CMIP5 climate sensitivity estimates. *Climate Dynamics*, *41*(11-12), 3339–3362. doi: 10.1007/s00382-013-1725-9

Wing, A. A., Reed, K. A., Satoh, M., Stevens, B., Bony, S., & Ohno, T. (2017). Radiative-convective equilibrium model intercomparison project. *Geoscientific Model Development*. doi: 10.5194/gmd-2017-213

Xue, Y., Li, Y., Li, Z., Lu, R., Gunshor, M., Moeller, S., … Schmit, T. (2020). Assessment of upper tropospheric water vapor monthly variation in reanalyses with near-global homogenized 6.5-μm radiances from geostationary satellites. *Journal of Geophysical Research: Atmospheres*, *125*(18). doi: 10.1029/2020JD032695

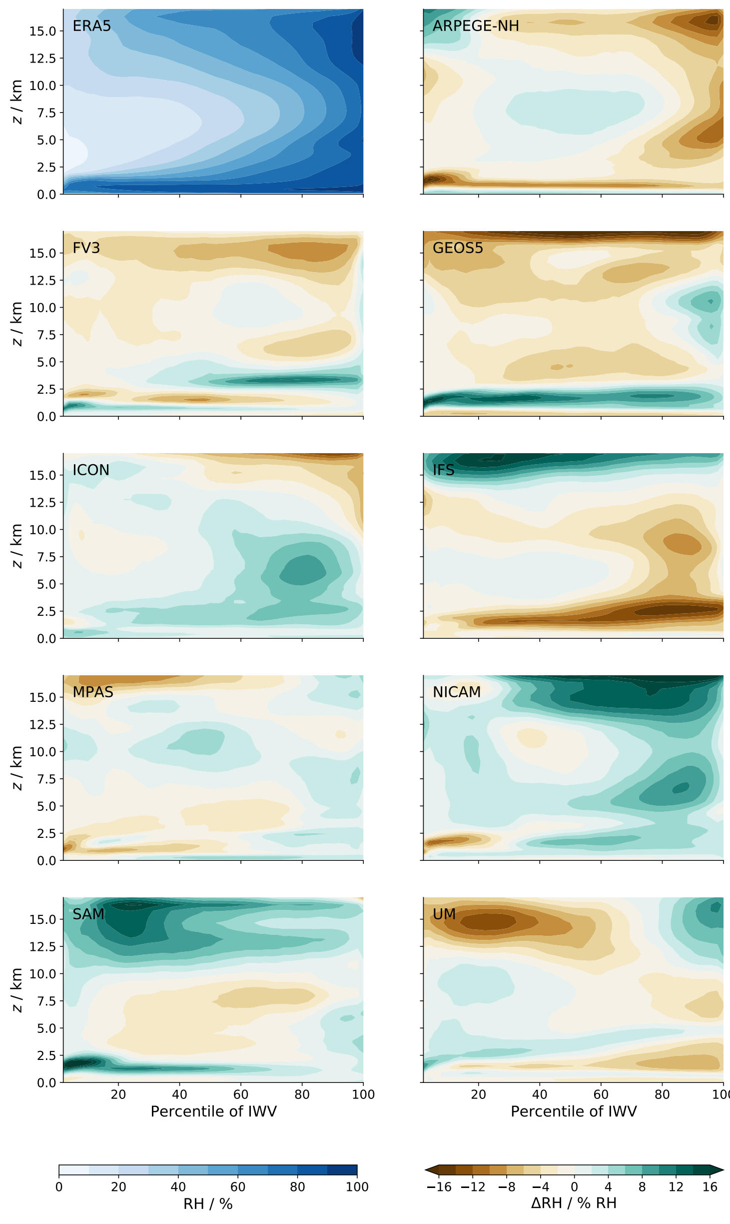

Figure A1. RH anomalies of DYAMOND models in moisture space. The upper left panel shows the ERA5 RH distribution in moisture space, remaining panels show the deviation from the ERA5 RH for each model.

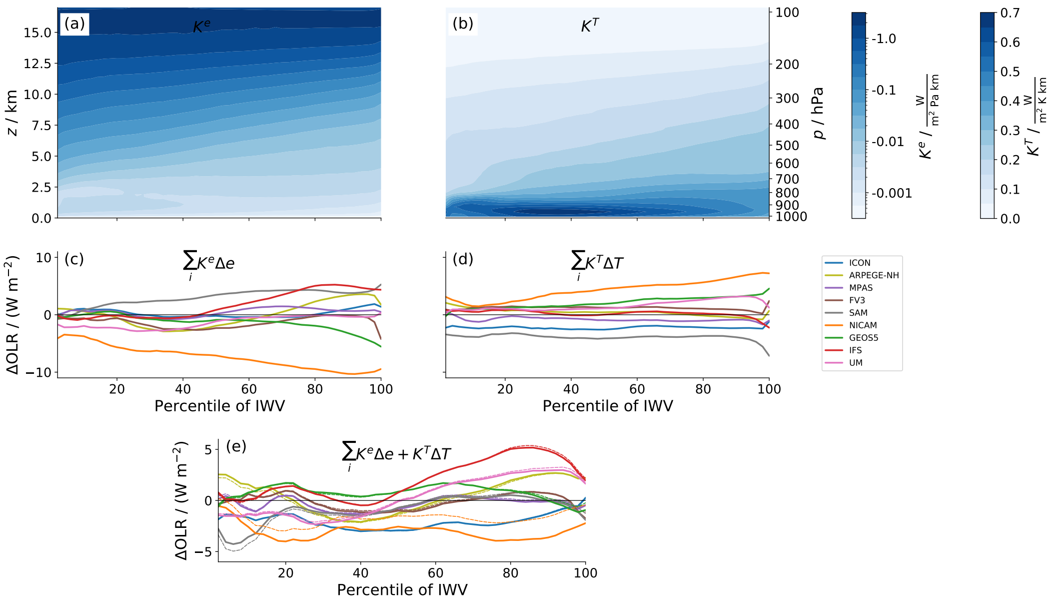

Figure B1. Clear-sky OLR anomalies in the DYAMOND models approximated with the kernel method. (a) Water vapour response kernel $\mathbf{K^e}$ showing the sensitivity of clear-sky OLR to a change of 1 Pa in water vapour pressure e in a 1 km layer. Note the logarithmic colour scale. (b) Temperature response kernel K^T showing the sensitivity of clear-sky OLR to a temperature change of 1 K in a 1 km layer. Also shown are OLR anomalies calculated (c) solely from anomalies in e and the respective kernel $\mathbf{K^e}$ and (d) solely from anomalies in T and $\mathbf{K^T}$. (e) shows OLR anomalies calculated from both kernels. True (directly calculated) OLR anomalies are shown as thin dashed lines for comparison.

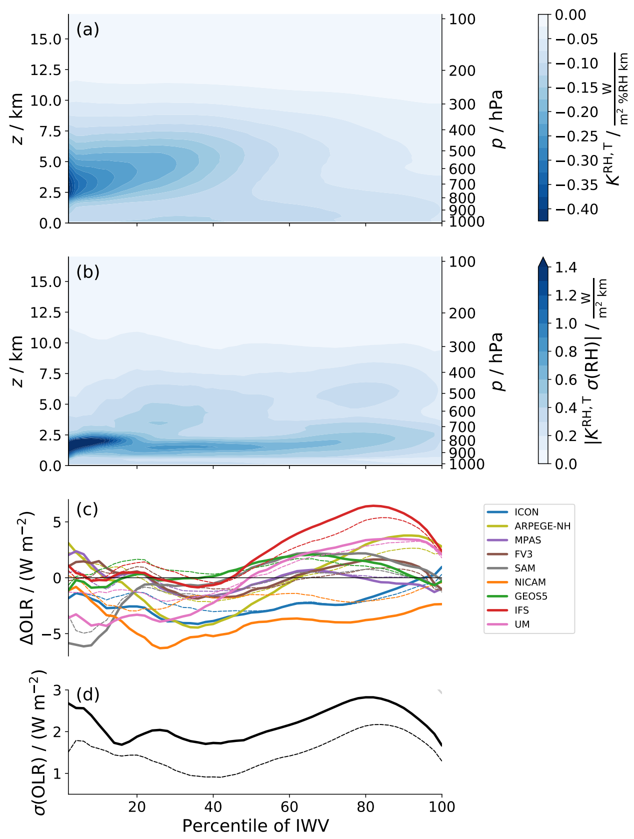

Figure B2. As Figure 8 but based on $\mathbf{K}^{\mathrm{RH},\mathbf{T}}$. Note that the colour scale in (a) and (b) is different from Figure 8 since $\mathbf{K}^{\mathrm{RH},\mathbf{T}}$ takes on more negative values than $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$.