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

Theresa Lang[1,1], Ann Kristin Naumann[2,2], Bjorn Stevens[2,2], and Stefan Alexander Buehler[1,1]

[1]Universität Hamburg
[2]Max Planck Institute for Meteorology

November 30, 2022

Abstract

Reducing the model spread in free-tropospheric relative humidity (RH) and its response to warming is a crucial step towards reducing the uncertainty in clear-sky climate sensitivity, a step that is hoped to be taken with recently developed global storm-resolving models (GSRMs). In this study we quantify the inter-model differences in tropical present-day RH across GSRMs, making use of DYAMOND, a first 40-day intercomparison. We find that the inter-model spread in tropical mean free-tropospheric RH is reduced compared to conventional atmospheric models, except from the the tropopause region and the transition to the boundary layer. We estimate the reduction to approximately 50-70% in the upper troposphere and 25-50% in the mid troposphere. However, the remaining RH differences still result in a spread of 1.2 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 clear-sky OLR: rather moist regimes at the transition from deep convective to subsidence regimes and very dry subsidence regimes. Particularly for these regimes a better understanding of the processes controlling the RH biases is needed.

Tropical free-tropospheric humidity differences and their effect on the clear-sky radiation budget in global storm-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:

- A 40-day comparison of storm-resolving models indicates that free-tropospheric relative humidity differs less than among conventional models
- The remaining relative humidity differences still cause a non-negligible (approximately 1.2 $\mathrm{W m^{-2}}$) spread in tropical mean clear-sky OLR
- Reducing humidity biases is most beneficial in the lower and mid free troposphere of dry subsidence regimes and near deep convective regimes

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

Abstract

Reducing the model spread in free-tropospheric relative humidity (RH) and its response to warming is a crucial step towards reducing the uncertainty in clear-sky climate sensitivity, a step that is hoped to be taken with recently developed global storm-resolving models (GSRMs). In this study we quantify the inter-model differences in tropical present-day RH across GSRMs, making use of DYAMOND, a first 40-day intercomparison. We find that the inter-model spread in tropical mean free-tropospheric RH is reduced compared to conventional atmospheric models, except from the the tropopause region and the transition to the boundary layer. We estimate the reduction to approximately 50-70% in the upper troposphere and 25-50% in the mid troposphere. However, the remaining RH differences still result in a 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 clear-sky OLR: rather moist regimes at the transition from deep convective to subsidence regimes and very dry subsidence regimes. Particularly for these regimes a better understanding of the processes controlling the RH biases is needed.

Plain Language Summary

Errors in the humidity and its change with global warming simulated by climate models limit our ability to predict how the climate system responds to an increase in greenhouse gas concentrations. In this study we investigate how large these humidity errors are in recently developed high-resolution models. We focus on the relative humidity, which measures the amount of moisture in the air compared to what air can hold at a given temperature. We find that the disagreement in the tropics is reduced compared to conventional climate models, but the relative humidity errors still have a considerable effect on the radiation budget. We also investigate in which regions of the tropics a further reduction of errors 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. Particularly for those regimes a better understanding of the processes that cause the model errors is needed.

1 Introduction

Free-tropospheric water vapor strongly impacts the Earth's outgoing longwave radiation (OLR) and therefore plays a key role in controlling the clear-sky response of the climate system to an increase in greenhouse gases. It is now widely accepted that this response is described by a warming and moistening of the atmosphere that is implied if the relative humidity (RH) and lapse rate were to depend on temperature alone, which corresponds to a warming at approximately constant RH (e.g. Held & Soden, 2000; Romps, 2014; Po-Chedley et al., 2019). This reduces the radiative response compared to a warming at constant absolute humidity, and can be described as a positive water-vapor–lapse-rate feedback. While general circulation models (GCMs) agree on this basic response (e.g. Soden & Held, 2006; Bony et al., 2006), there is still an appreciable inter-model spread in the magnitude of the water-vapor–lapse-rate feedback. This spread, which primarily originates from the tropics, contributes a non-negligible (about 30%) uncertainty to the climate sensitivity (Vial et al., 2013).

The RH is an important detail. Even small deviations from its assumed constancy with warming have a strong impact on the radiative response. RH changes alter the radiative compensation between water-vapor and lapse-rate feedback in the saturated regions of the emission spectrum (Bony et al., 2006) and differences in the RH response control the spread in tropical water-vapor–lapse-rate feedback across GCMs (Vial et al., 2013; Po-Chedley et al., 2018). Even if RH does not change with warming, the RH profile in the present climate may influence the feedback. While a correlation between global mean present-day humidity and water vapor feedback has not been found for GCMs (John & Soden, 2007), Bourdin et al. (2021) have argued that especially at warmer, tropical temperatures the rapid closing of the atmospheric window by water vapor continuum absorption makes the feedback dependent on the RH profile. There are other reasons to care about present-day free-tropospheric RH (e.g. Derbyshire et al., 2004; Luo & Rossow, 2004; Stevens et al., 2017), but independent of whether these (or the proposed direct effect of present-day RH on the feedback) end up being important, confidence in an ability of models to correctly represent the present-day RH is essential for building trust in model-based estimates of the subtle changes in RH under warming that influence the water vapor feedback.

Sherwood et al. (2010) found that certain aspects of the tropical RH distribution show signs of convergence in GCMs once horizontal resolutions fall below about 100km. It is also known from previous studies that free-tropospheric RH is primarily controlled by the circulation on scales resolved by typical GCMs, and parameterized processes like convection only matter by influencing the circulation (e.g. Sherwood, 1996; Pierrehumbert & Roca, 1998; Dessler & Sherwood, 2000). On the one hand, the convergence of RH in GCMs with different convective parameterizations might indicate that convective processes play a minor role in affecting the circulation. On the other hand, for simulations on an aquaplanet Retsch et al. (2019) found that allowing convection to be resolved explicitly has a larger impact on free-tropospheric RH than increasing resolution in simulations with parameterized convection. This suggests that the circulation changes more significantly once convection is resolved explicitly and calls into question whether the RH in GCMs converges for physical reasons.

A milestone in climate modelling has been made with the emergence of global storm-resolving models (GSRMs; Satoh et al., 2019), also called global cloud-resolving or convection-permitting models. While the development of the first GSRM already goes back more than 15 years (Tomita et al., 2005), only recently the increase in computational capacities has allowed several modelling groups to follow, enabling first intercomparisons. GSRMs 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, which is hoped to eradicate some long-standing biases (e.g. Miura et al., 2007; Stevens et al., 2020). Whether the spread in free-tropospheric RH is reduced in GSRMs is, however, not obvious. This depends on how strongly the behavior of convection depends on model formulation. If this dependence is weak, RH differences should be small among GSRMs. However, there are also reasons to expect the opposite. Bourdin et al. (2021) found that RH differences across cloud-resolving models in radiative-convective equilibrium (RCE) are substantially larger than across GCMs. The large spread in RCE models is likely related to different degrees of convective organization (Becker & Wing, 2020). Although these differences are expected to be smaller in simulations with more realistic setups, in which large-scale circulations impose constraints on convective organization (Wing et al., 2020), they likely still play a role. Therefore, it cannot be ruled out that the RH spread across GSRMs is similar or even larger than across GCMs.

In this study we quantify differences in tropical free-tropospheric RH across GSRMs for the first time, making use of the model intercomparison DYnamics of the Atmospheric general circulation Modeled On Nonhydrostatic Domains (DYAMOND Stevens et al., 2019). To assess how relevant the RH differences are from a radiative point of view, we translate them into differences in clear-sky outgoing longwave radiation (OLR) using a radiative transfer scheme. The latter is also used to compute radiative kernels, which allow us to identify those regions in the tropical atmosphere, in which a future reduction of RH differences would be 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. On the one hand, humidity fields in moisture space are highly aggregated, which ensures robust statistics. On the other hand, the moisture space representation allows us to distinguish between different dynamic regimes of the tropics, which is useful for identifying regions of large inter-model spread as well as for the OLR calculations. 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 RH differences in the tropical mean and in moisture space. The impact of the RH 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 among the models. The models were not specifically calibrated for the DYAMOND simulations. Some models even ran for the first time in this configuration and at storm-resolving resolutions.

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 vertical velocity 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.

The size of the model output represents a challenge for the analysis. 30 days of one 3-hourly 3D field (corresponding to 240 timesteps) on the native model grid covering the tropics have a size on the order of 2 TB. For nine models and four variables this adds up to more than 60 TB. Developing strategies for dealing effectively with the massive amounts of data produced by GSRMs is one of the purposes of DYAMOND. Our approach is the following: In a first step all fields are horizontally interpolated from each model's native 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 remapping reduces the data volume by about a factor of ten without noticeable loss of information in the region of interest. In a second step we perform a subsampling of grid points. From each of the 240 output timesteps about 42,000 oceanic profiles are selected randomly, resulting in a total of 10 million selected profiles for each model. This reduces the amount of data by another factor of 100. We estimated the sampling uncertainty by repeating the random sampling several times for the same model. For tropical mean RH, the quantity we focus on, the sampling uncertainty is about 0.01% RH and hence two magnitudes smaller than inter-model differences, which are on the order of 1% RH (Section 3.1). In the same manner we estimated the sampling uncertainty for each block in moisture space (Section 3.2) to be at least one order of magnitude smaller than the inter-model spread in the respective block. Hence, the random subsampling of profiles introduces only a small error, but reduces the data volume to 0.1% of its original size. This result shows that although GSRMs work with tremendous data volumes, most of the information is necessary for predicting their dynamic evolution, and for many analyses there exists considerable opportunities to compress their output with relatively little loss of information.

We exclude land areas to avoid complications from topography and more strongly varying boundary layer depths and hence to simplify the interpretation. The inhomogeneity of land regions would also colour our analysis in moisture space. Vertically integrated water vapor (IWV), which is used to span moisture space (Section 3.2), is strongly influenced by local surface characteristics over land. It can be very low in regions with little soil moisture or in regions with high elevation. Consequently, if moisture space was spanned from both oceanic and continental grid points, profiles associated with very different regimes would be mixed in the same IWV blocks. Therefore, we focus on the more homogeneous ocean regions.

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 water vapor 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 RH differences in DYAMOND models

In this section we quantify the differences in free-tropospheric RH in the DYAMOND models, first in the tropical mean and subsequently in moisture space.

3.1 Tropical mean

Since the focus of this study is on the radiative impact of humidity differences we concentrate on relative humidity (RH) rather than absolute humidity (measured by q). The atmospheric temperature and water vapor concentration are decisive parameters for clear-sky radiative transfer. The RH is a valuable proxy that links their competing effects on longwave emission. This will be discussed in more detail in the second part of this paper. Another reason to look at RH is that it is RH rather than q that is effectively constrained by model processes (in particular, condensation and evaporation). Therefore, any model errors in temperature are expected to alter q but not necessarily RH.

RH is calculated for each of the randomly selected profiles and their associated values of q, p and T as $\mathrm{RH} = \frac{e}{e_s(T)}$, where e is the water vapor 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 and 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). It should be noted here that the RH computed in this way can deviate from the RH calculated internally in the microphysics schemes of the models because they use different methods to compute RH above the freezing level. The deviations are relevant when the relation between RH and clouds or precipitation is investigated. However, as explained above our focus is on the radiative impact of the humidity differences. We regard RH primarily as a quantity that links temperature and absolute humidity, which are the quantities that ultimately enter the models' radiation schemes. Therefore, it is reasonable to compare RH computed in a uniform way for all models.

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 $\sigma(\mathrm{RH})$ (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 RH 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.

To the extent one thinks of RH anomalies as linking q and T anomalies, it is informative to consider q and T separately. In the DYAMOND models, T anomalies are smallest in the lower troposphere, 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). At lower levels, where T anomalies are small, q and RH anomalies are correlated (Figure 1b, Figure 2d). In the upper troposphere, where T anomalies are large, T and q anomalies are correlated (Figure 2b,d), consistent with the idea that model errors in T cause errors in q. Although RH anomalies are also large there (Figure 1), they play a minor role in determining whether a model's q is small or large as compared to another model's q.

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 water vapor 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 mostly systematic model biases rather than a result of poorly sampled internal variability. The region where the inter-model differences are expected to be colored most strongly by internal variability is the upper troposphere, where the inter-model spread is only two to three times larger than the estimated internal variability.

Another potential limitation arises from the common initialization of the models, which might constrain the RH profiles even after the first ten days of the simulation that were excluded (Section 2.2). To test this, we divided the analyzed 30-day period into three consecutive 10-day periods and repeated the spread analysis. We did not find a systematic increase of the inter-model spread over time, except from the altitude region above 14 km. For a second analysis we made use of a coupled atmosphere-ocean simulation performed with the ICON model at storm-resolving resolution (5 km grid spacing). The simulation was run for two years, starting on 20th January 2020. The length of the simulation allows us to examine how the RH profile evolves after the first 40 days. In Figure 3 we compare tropical mean RH profiles for February 2020 and February 2021. February 2020 corresponds to days 13 to 40 after the initialization and is hence comparable to the time period we analyze in the DYAMOND simulations. If the RH profile was still in the transition from the initial conditions during that month, we would expect it to be very different one year later. However, the RH differences between February 2020 and February 2021 are small compared to the inter-model differences (cf. Figure 1). Throughout the lower and mid troposphere, the difference is smaller than 1% RH. The largest differences of up to 3% RH occur in the upper troposphere above 12 km. It has to be kept in mind that SST changes from February 2020 to February 2021 in the coupled simulation, so the RH differences we find are most likely related to SST changes rather than to constraints from initialization in February 2020. The size of the differences and the increase in the upper troposphere are in accordance to what we found for the inter-annual variations in ERA5 (Figure 1c). It is very unlikely that the RH in February 2020 was still in its transition from initialization, but SST and/or model drift changed in a way to keep RH almost constant in February 2021. Hence, both analyses indicate that the transition from the initial conditions is already largely completed after the first ten days. The upper troposphere (above 12 km) might be an exception, but as we will see in Section 4 the RH differences in this region do not significantly affect the clear-sky radiation budget.

To examine how the RH spread in DYAMOND compares to that in conventional, coarser atmospheric GCMs, we compare the DYAMOND ensemble to 29 GCMs that participated in the Atmospheric Model Intercomparison Project (AMIP) experiments of the Coupled Model Intercomparison Project phase five (CMIP5) (Taylor et al., 2012). The AMIP simulations have a total length of 30 years (1979-2008) and were run with prescribed (identical) SST. An exact quantitative comparison of the RH spread in GSRMs and GCMs will not be possible until longer, multi-year storm-resolving simulations are available. Nevertheless, a comparison to the AMIP GCMs is valuable to put the DYAMOND spread into perspective. The inter-model spread in AMIP is quantified both based on 30-year averages and based on monthly averages of RH. This allows us to estimate how much the inter-model spread in a single month can differ from the spread on climatological timescales. The inter-model standard deviation of 30-year mean RH is denoted by the black dashed line in Figure 1c. It lies within the range of monthly standard deviations, which is shown as gray shading. In most parts of the free troposphere, the most extreme monthly standard deviations differ between 5-25% from the 30-year value. Only in the tropopause region the deviations are larger (up to 40%). Overall, the AMIP experiment confirms that the inter-model spread in a single month provides a good first estimate of the inter-model spread on climatological timescales. However, the the variability in the monthly standard-deviation should be kept in mind when the (monthly) DYAMOND spread is compared to the (climatological) AMIP spread in the following.

The inter-model spread in DYAMOND is smaller than the spread in AMIP throughout most of the free troposphere. The largest reduction is found between 8 km and 14 km altitude, where the RH spread in DYAMOND is reduced by approximately 50-70% compared to AMIP. At lower altitudes, between 3 km and 8 km altitude, the DYAMOND spread is smaller by approximately 25-50%. The lower free troposphere is an exception: the peak in σ(RH) at the top of the BL is less pronounced in CMIP5 AMIP than in DYAMOND, indicating that variations in the depth of the BL are smaller in the AMIP models. However, part of the smaller spread in the AMIP models can be explained by the fact that the hydrolapse in these models is generally less steep, which is evident from the AMIP 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.

As mentioned in Section 1, Sherwood et al. (2010) found that certain aspects of the RH distribution converge in GCMs once horizontal grid spacings fall below a certain scale. A question arising from this is whether the agreement across GSRMs is better than across the CMIP5 AMIP models with rather high resolutions. To test this we repeated the spread analysis for only those nine AMIP models with grid resolutions exceeding T85 (128x256 grid points), corresponding to the scale suggested by Sherwood et al. (2010). While the RH spread across these high-resolution GCMs is somewhat reduced in the upper and lower troposphere, the spread in the mid troposphere seems to be unaffected (not shown). As we will show in Section 4.4, it is particularly the spread in the mid troposphere that matters for the outgoing longwave radiation. Hence, there is still a valuable improvement in GSRMs compared to the high-resolution GCMs.

An additional series of DYAMOND runs with the ICON model allowed us to investigate how RH changes with increasing horizontal resolution beyond the convergence scale suggested by Sherwood et al. (2010). We compared tropical mean (ocean only) RH from runs at 80 km, 40 km and 20 km grid spacing with parameterized convection as well as runs at 20 km, 10 km, 5 km and 2.5 km grid spacing with explicit convection (not shown). In the parameterized runs RH hardly changes with increasing horizontal resolution. RH strongly depends on resolution for the explicit runs at 20 km and 10 km, for which using explicit convection might not be adequate. At 5 km grid spacing RH has converged. In some altitude regions, particularly in the mid troposphere, the RH difference between the converged explicit runs and the parameterized runs is significantly larger than the differences between the parameterized runs at different resolutions. These findings suggest that resolving convection impacts RH although it seemed to have already converged at coarser resolutions when convection was parameterized.

In summary, despite the shortness of the DYAMOND simulations we can say with a high degree of certainty that the spread in free-tropospheric RH in the DYAMOND GSRMs is reduced compared to the AMIP GCMs throughout most of the free troposphere, except from the region at the transition to the BL and the tropopause region. We estimate the reduction to approximately 50-70% in the upper troposphere (8-14 km) and 25-50% in the mid troposphere (3-8 km). For an exact quantification longer storm-resolving simulations are required. The reduction in the spread is even more remarkable considering that the DYAMOND models were not specifically calibrated 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 RH differences still have a non-negligible impact on the clear-sky radiation budget.

3.2 Moisture space

To distinguish between different dynamic regimes of the tropics, namely subsidence and deep convective regimes, which are not necessarily co-located in different models, we compare RH 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 vapor (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 4). 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 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 prob-

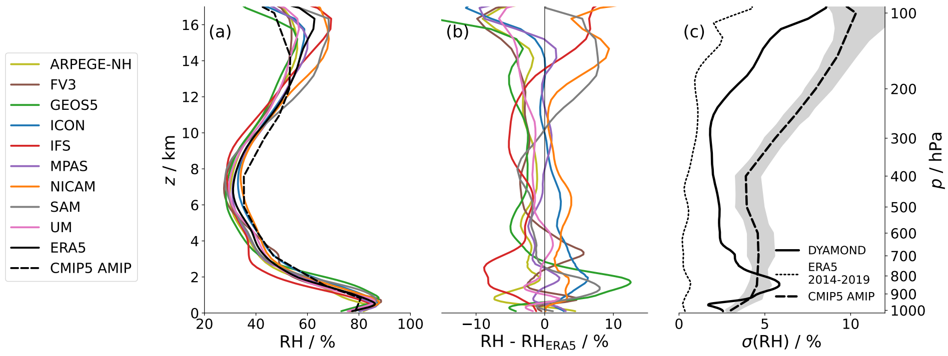

Figure 1. Tropical mean RH profiles and inter-model spread in the DYAMOND ensemble. (a) Tropical mean vertical profiles of RH over ocean regions from all DYAMOND models (colours), the ERA5 reanalysis (black solid) and the CMIP5 AMIP 30-year multi-model mean (black dashed). (b) Vertical RH profiles for the DYAMOND models shown as deviation from the ERA5 profile. (c) Inter-model standard deviation of tropical mean RH in DYAMOND (solid line). For comparison, the inter-annual RH spread in five years of ERA5 (2014-2019; dotted line) as well as the inter-model spread of the 30-year mean RH in the CMIP5 AMIP ensemble (dashed line) are shown. Grey shading indicates the range of inter-model standard deviations in individual months of the AMIP experiment.

lem 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. This block-averaging results in an x-axis that is linear in the percentile of IWV. Due to the non-uniform IWV distributions (Figure 4) block-averaged IWV itself increases non-linearly as a function of the IWV percentile. This is also visible in the multi-model mean (black line in Figure 5d), albeit very weakly: In the driest and moistest percentiles, respectively, the increase in IWV is steeper than in the intermediate percentiles. Note that 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.

SST increases from about 292 K in low IWV percentiles to about 302 K in high percentiles (Figure 5d). 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 5c) 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 rainfall in the Indo-Pacific Warm Pool and the Intertropical Convergenze Zone (ITCZ), where deep convection is concentrated. Note that block-averaged vertical velocities take on values up to 13 cm s^{-1} in the deep convective regimes, but the color map in Figure 5c is truncated at 1.2 cm s^{-1}. The drier blocks correspond to trade wind regimes. There,

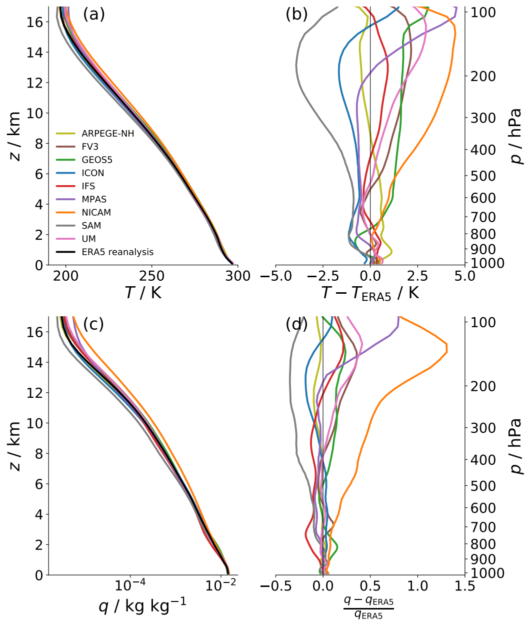

Figure 2. Tropical mean vertical profiles of temperature T and specific humidity q over ocean regions from all DYAMOND models. Vertical profiles of T (a, b) and q (c, d) are shown as absolute values together with the ERA5 profiles (a, c) and as deviation from the ERA5 profiles (b, d). Deviations in q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

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 increasing amount of high-level clouds from dry to moist regimes is also reflected by a sharp decrease in all-sky OLR in the moist blocks (Figure 5d).

The largest RH values are found in the BL (5a), 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. Throughout the tropics, particularly in the subsidence regions, the free-tropospheric RH profile takes on a typical C-shape, which is known from observations (e.g. Jensen et al., 1999; Vömel et al., 2002) and GCMs (Sherwood et al., 2010). With a simple analytical model Romps (2014) showed that this shape of the RH profile can be understood from the balance between moistening by detrainment of saturated air from convective regions and drying by subsidence. As the temperature lapse rate increases with height, the reduction in RH for a given amount of subsidence also increases with height. This increase in subsidence drying, together with a decrease in convective moistening, explains why RH decreases with height in the lower free troposphere. In the upper troposphere, however, convective moistening dominates and causes RH to maximize at the tropopause. A plateau in

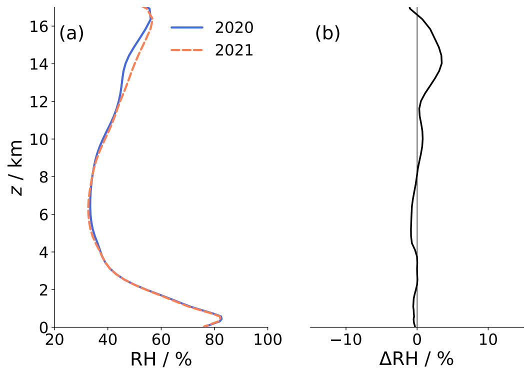

Figure 3. Comparison of RH in two subsequent Februaries of a coupled atmosphere-ocean simulation with the ICON model at storm-resolving resolution (5 km). (a) Tropical mean (ocean only) RH in February 2020 (blue) and February 2021 (orange). February 2020 corresponds to days 13 to 40 after initialization, which is comparable to the analyzed DYAMOND period. (b) RH difference between February 2020 and February 2021.

RH is apparent near the freezing level at around 5 km particularly in the high IWV percentiles. 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 5b. 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 5a). 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 5c). The large spread in this regime might be related to model differences in convective behavior. 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.

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 ($\mathrm{OLR_c}$) using a radiative transfer model. The differences are analyzed in moisture space to determine how much different tropical moisture regimes contribute to the inter-model spread in tropical mean $\mathrm{OLR_c}$. Furthermore, we use radiative kernels to examine in which altitude regions RH differences have the strongest impact on $\mathrm{OLR_c}$. This allows us to identify the regions of the tropical troposphere in which a further reduction of RH differences would be most beneficial.

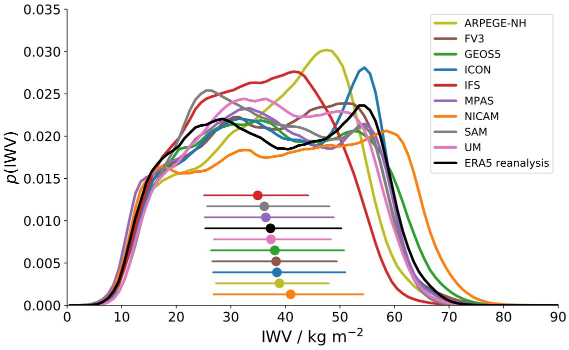

Figure 4. Probability density function of integrated water vapor (IWV) over tropical ocean regions in the DYAMOND models and ERA5. Percentiles of each model's IWV distribution are shown below the curves: Coloured circles indicate the median, horizontal bars range from the 25th to the 75th percentile.

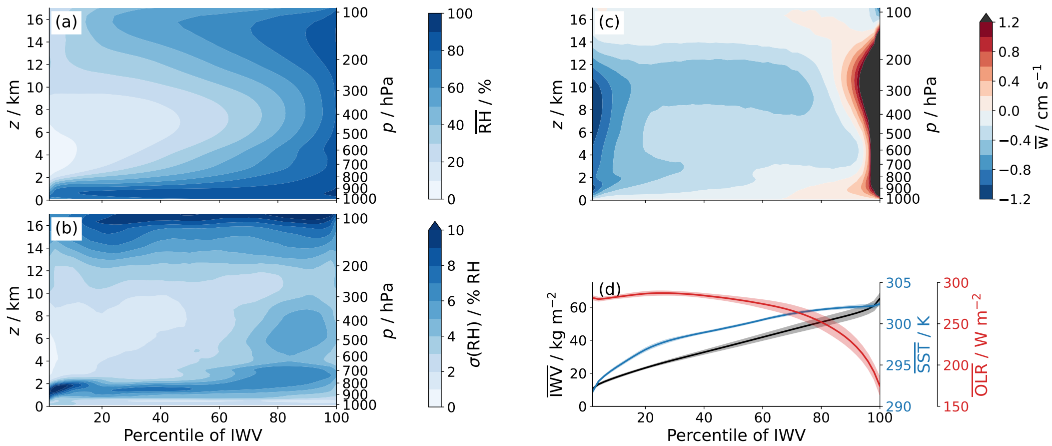

Figure 5. Distributions of different block-averaged quantities in moisture space: (a) multi-model mean RH, (b) multi-model standard deviation of RH, (c) multi-model mean vertical velocity and (d) multi-model mean IWV (black), SST (blue) and all-sky OLR (red). Note that the color map for vertical velocity in (c) is truncated at $1.2\,\mathrm{cm\,s^{-1}}$ and any larger values (up to $13\,\mathrm{cm\,s^{-1}}$ in the highest IWV block) are displayed in black. For the quantities in (d) the inter-model standard deviation is denoted by shaded areas around the multi-model mean values.

Fundamentally, clear-sky OLR is determined by surface temperature as well as atmospheric temperature and greenhouse gas concentrations. For the $\mathrm{OLR_c}$ anomalies in the DYAMOND models we expect that anomalies in the surface temperature play a minor role, since SST is prescribed and its distributions in moisture space is very similar among models (Figure 5). Furthermore, compared to model differences in water vapor we expect differences in other greenhouse gasses to have a small effect on $\mathrm{OLR_c}$. Therefore, we fix the concentrations of other greenhouse gasses in our radiative transfer simulations. Thus, we assume that $\mathrm{OLR_c}$ anomalies in the DYAMOND models 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. 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). Note that not all of the DYAMOND models employ RRTMG as their native radiation scheme. Differences in the radiation codes can cause errors on the order of 2 $\mathrm{Wm^{-2}}$ in the models' internally calculated clear-sky OLR (Pincus et al., 2015). By using the same radiation scheme for each model for our offline calculations we neglect this error source, but instead focus solely on the effect of RH differences on clear-sky OLR.

$\mathrm{OLR_c}$ is calculated based on the block-averaged profiles of pressure, temperature, and specific humidity in moisture space (Section 3.2). We found that calculating $\mathrm{OLR_c}$ from block-averaged profiles generally introduces a small negative error compared to $\mathrm{OLR_c}$ calculated based on individual profiles. OLR is often thought to increase linearly with temperature, and does, increasingly so, as temperatures are reduced below their tropical mean (e.g. Koll & Cronin, 2018). Within the tropics, where temperature fluctuations are small, variability in clear-sky OLR is dominated by RH changes (e.g. John et al., 2006). Due to the approximately logarithmic dependence of $\mathrm{OLR_c}$ on RH, averaging decreases $\mathrm{OLR_c}$ (Pierrehumbert et al., 2007). However, the resulting bias is very similar for all models, so that the effect on inter-model differences in $\mathrm{OLR_c}$ is negligible.

To characterize the surface we use model output of surface pressure and the prescribed SST fields and select the same points as for the 3D data (Section 2.2). The surface emissivity is assumed to be 1. For other gasses than water vapor 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 $\mathrm{OLR_c}$ is defined as the longwave upward clear-sky radiative flux at this level. Due to this definition the inter-model differences in $\mathrm{OLR_c}$ only reflect T and q differences in the troposphere, potential differences in the stratosphere are ignored. Note that due to the missing stratosphere the absolute value of the $\mathrm{OLR_c}$ defined at 20 km has a positive offset compared to the "true" $\mathrm{OLR_c}$ 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 $\mathrm{OLR_c}$ differs by more than 4 $\mathrm{Wm^{-2}}$ between the two most extreme models IFS and ICON (Figure 6a). The multi-model standard deviation in tropical mean $\mathrm{OLR_c}$ is 1.2 $\mathrm{Wm^{-2}}$. This is small compared to cloud radiative effects, but still a third of the estimated radiative forcing due to a doubling of CO_2 (Collins et al., 2013). In some models, e.g. UM and ARPEGE-NH, both positive and negative 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 6b): 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 5b). 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 spread in $\mathrm{OLR_c}$ maximizes in these two regimes and which altitude regions in the troposphere contribute most.

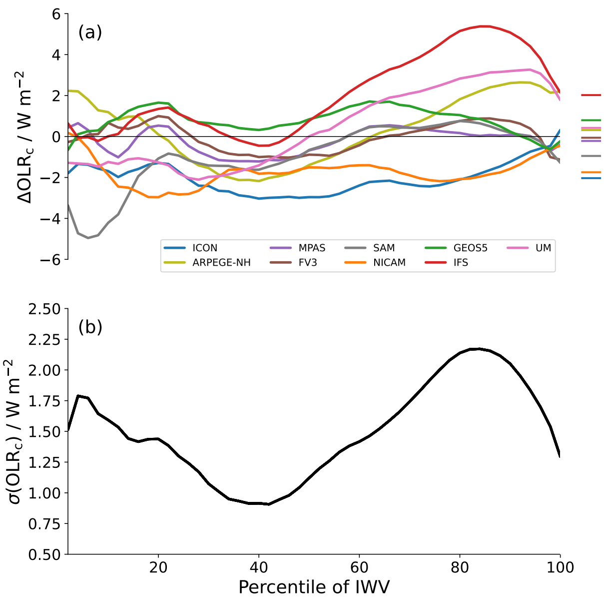

Figure 6. Inter-model differences in clear-sky OLR in moisture space. (a) Anomalies in clear-sky OLR for each model, defined as the deviation from the ERA5 value and (b) inter-model standard deviation of clear-sky OLR.

4.3 Radiative kernels

To examine how different altitude regions in moisture space contribute to the spread in tropical mean $\mathrm{OLR_c}$, for each of the 50 blocks in moisture space we decompose each model's $\mathrm{OLR_c}$ 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 clear-sky OLR anomaly $\Delta\mathrm{OLR_c}$ can be written as:

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

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

$$K_i^x = \frac{\partial \mathrm{OLR_c}}{\partial \mathrm{x}_i}. \tag{2}$$

The first approximation in Equation 1 assumes that anomalies in $\mathrm{OLR_c}$ are primarily caused by anomalies in atmospheric e and T, 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 from perturbations above. Despite these assumptions the kernels $\mathbf{K^e}$ and $\mathbf{K^T}$ can be used to approximate the $\mathrm{OLR_c}$ anomalies of the DYAMOND models with good accuracy, which is shown in Figure B1 in Appendix B. The computation of the kernels is also described in Appendix B.

Perturbations in e and T have opposite effects on $\mathrm{OLR_c}$, which is evident from the different signs of the respective kernels (Figure B1). At constant RH perturbations in e and T are positively correlated, so their effects on $\mathrm{OLR_c}$ compensate to some degree. It is well known that in the water vapor bands, the spectral regions at which the water vapor 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 behavior is often referred to as "Simpsonian", as it has been recognized since the early work of Simpson (1928). Therefore, it can be assumed that anomalies in $\mathrm{OLR_c}$ in the DYAMOND models are primarily determined by RH anomalies. This corresponds to the second approximation in Equation 1.

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:

$$\begin{aligned} K_i^{\mathrm{RH},e} &= \frac{\partial \mathrm{OLR_c}}{\partial \mathrm{RH}_i}\bigg|_{T=\mathrm{const.}} &&= e_s\, K_i^e \\ K_i^{\mathrm{RH},T} &= \frac{\partial \mathrm{OLR_c}}{\partial \mathrm{RH}_i}\bigg|_{e=\mathrm{const.}} &&= -\frac{e_s}{\mathrm{RH}}\left(\frac{de_s}{dT}\right)^{-1} K_i^T. \end{aligned} \tag{3}$$

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 vapor 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 $\mathrm{OLR_c}$ response to a given change in RH is independent of whether this change is produced by a change in e or in T.

$\mathrm{OLR_c}$ anomalies approximated using $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ (Figure 7c) are more accurate than those approximated using $\mathbf{K}^{\mathrm{RH},\mathbf{T}}$ (Figure B2c). Therefore, for the further analysis we concentrate on $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$. Overall, $\mathrm{OLR_c}$ anomalies approximated from RH anomalies agree well with true (directly calculated) $\mathrm{OLR_c}$ anomalies (Figure 7c) and the inter-model standard deviation $\sigma(\mathrm{OLR_c})$ is well reproduced (Figure 7d). In Appendix B we elaborate more on the accuracy of the approximation for individual models as well as on the differences between $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ and $\mathbf{K}^{\mathrm{RH},\mathbf{T}}$.

4.4 Relative importance of different altitude regions

The impact of RH anomalies for the radiation budget is determined by the magnitude of the RH anomalies and the sensitivity of $\mathrm{OLR_c}$ to a given perturbation in RH. The latter is described by the radiative kernel $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ (Equation 1). $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ is negative throughout the tropical troposphere (Figure 7a), indicating that an increase in RH leads to a decrease in $\mathrm{OLR_c}$. Its absolute value is largest in the mid troposphere in the dry subsidence regimes.

The overall distribution of the kernel can be understood based on the concept of an effective emission height for each wavenumber ν, corresponding to the level at which the optical depth τ_ν reaches unity (e.g. Petty, 2006). A water vapor perturbation will generally have a strong impact on OLR if it is applied near or above a level for which $\tau_\nu \approx 1$ in a large portion of the water vapor bands. Ultimately, the vertical distribution of $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ is determined by the distribution of effective emission heights. The distribution of effective emission heights depends on the distribution of spectral absorption coefficients and is generally broad (e.g. Clough et al., 1992; Jeevanjee & Fueglistaler, 2020), which is why $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ is significant throughout the troposphere. However, above a certain level (around 200 hPa) the emission from water vapor rapidly declines, which is well known from studies of radiative cooling (e.g. Hartmann & Larson, 2002). Due to the strong dependence of water vapor concentrations on temperature through Clausius-Clapeyron, the amount of water vapor at these upper levels is so small that even at the line centers τ_ν barely reaches unity. The emission to space also declines at the lowest levels, although water vapor is abundant, because there is only a limited part of the spectrum (on the wings of lines and very weak lines), where radiation can escape to space without being re-absorbed at upper levels. This "masking" by the optically thick atmosphere above increases with increasing IWV, which is why for a given altitude level the absolute value of $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ decreases towards moist regimes.

Note that in general the distribution of a water vapor kernel is very sensitive to how water vapor is perturbed (Held & Soden, 2000). We perturb RH by a constant value, similar to Spencer and Braswell (1997) or Allan et al. (1999). In this case the perturbation in e is proportional to e_s (Equation 3). Hence, it decreases with altitude, but is approximately constant throughout moisture space. Other studies apply equal fractional perturbations in e (Shine & Sinha, 1991) or keep RH constant under a uniform temperature perturbation (Held & Soden, 2000; Soden et al., 2008). In both cases the perturbation in e is proportional to e itself, resulting in a stronger weighting of moist compared to dry regimes.

In low IWV percentiles $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ peaks at an altitude of around 6 km. The peak weakens from dry to moist regimes for the reasons named above. A very similar behavior 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, however, 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 $\mathrm{OLR_c}$ as RH increases in the lower troposphere. In contrast to absorption in the water vapor bands, continuum absorption scales with the square of the water vapor pressure and therefore becomes relatively more important for high humidity base states.

The product of the RH response kernel $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ and the RH inter-model standard deviation $\sigma(\mathrm{RH})$ (Figure 7b) 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_c is not particularly sensitive to RH perturbations there (Figure 7a), but the inter-model differences in RH are large (Figure 5b) because the models differ in the depth of the BL. RH differences in a broad layer in the mid troposphere also significantly affect OLR_c. 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 5b) and a decreasing sensitivity of OLR_c from dry to moist regimes (Figure 7a). 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_c generally extends to higher altitudes in the dry regimes than in the moist regimes, which is again a consequence of the stronger masking effect in moist regimes. Due to the low sensitivity of OLR_c to RH perturbations in the upper troposphere (above about 10–12 km) the large inter-model RH differences there (Figure 5b) have virtually no effect on OLR_c.

Not considering clouds has an effect on the response kernels. Particularly high clouds are important, because they mask some of the effect of T and q 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). In these regimes we would expect the sensitivity of OLR_c to RH perturbations to decrease, particularly in levels below the clouds, which are most abundant at around 8-12 km height. This would 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 7d, does not yield the inter-model standard deviation in OLR_c, 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_c 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 7d and evident from Figure A1. In fact, it is only due to these compensating effects that dry regimes contribute less to tropical mean differences in clear-sky OLR than moist regimes.

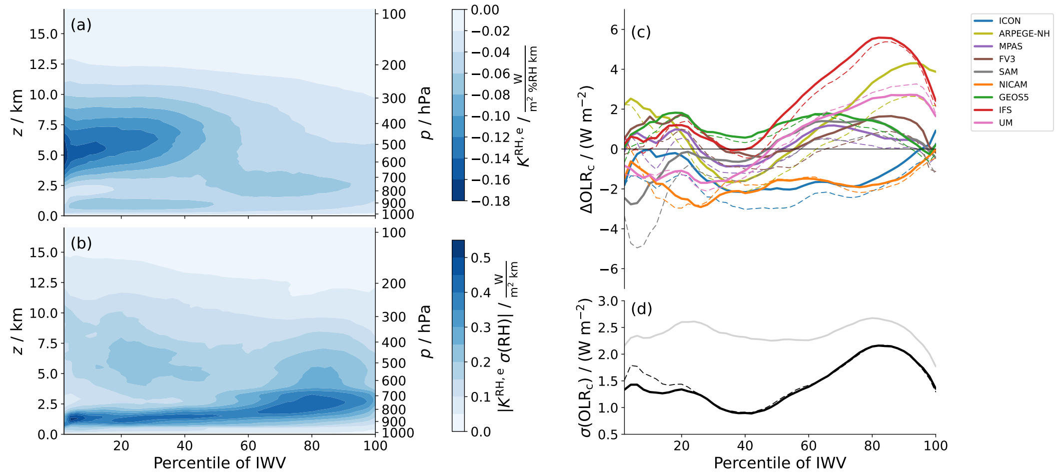

Figure 7. Impact of RH differences on clear-sky OLR in moisture space. (a) RH response kernel $\mathbf{K}^{\mathrm{RH,e}}$ showing the sensitivity of clear-sky OLR to a 1% RH change in a 1 km layer under constant temperature for 50 blocks in moisture space, (b) inter-model standard deviation σ(RH) weighted with $\mathbf{K}^{\mathrm{RH,e}}$, (c) Clear-sky 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 clear-sky OLR. Thin dashed lines in (c) and (d) correspond to "true" clear-sky OLR calculated directly from temperature and specific humidity profiles (same as in Figure 6). 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, which took part in DYAMOND, a first 40-day intercomparison 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).

A justified question that arises is how much one can learn about climatological RH biases from an intercomparison as short as 40 days. To address some major concerns associated with the shortness of the DYAMOND simulations, we performed additional analysis based on longer-term data sets. One potential limitation is that the models' RH might still be constrained by the common initial conditions. However, both a first two-year storm-resolving simulation with the ICON model as well as the evolution of the inter-model RH spread within the analyzed 30-day period suggest that the transition from the initial conditions is largely completed after the excluded ten-day spinup period. Another concern is that the RH biases identified in the analyzed 30-day period might result mainly from a poor sampling of internal variability. However, the DYAMOND inter-model spread in RH is significantly larger than what would be expected from internal variability, which was estimated from five years of ERA5 reanalysis data. This suggests that the inter-model differences we find in DYAMOND mostly represent systematic model biases. This applies least to the upper troposphere (above 12 km), where natural variability is comparably large. In accordance with that, the inter-model RH spread in each individual month of the CMIP5 AMIP intercomparison is within a 25% range of the spread in 30-year mean RH, only in the upper troposphere deviations are larger. We conclude from these results that in a large part of the free-troposphere one month of intercomparison already provides a good first estimate for climatological RH biases.

The comparison to the CMIP5 AMIP ensemble also shows that the inter-model spread in tropical mean RH in DYAMOND is reduced throughout the free troposphere, except from the transition to the boundary layer and the tropopause region. This indicates that free-tropospheric RH and hence clear-sky OLR are better constrained in GSRMs than in GCMs. Based on this first month of intercomparison we estimate the reduction to approximately 50-70% in the upper troposphere (8-14 km) and 25-50% in the mid troposphere (3-8 km). For an exact quantification longer storm-resolving simulations will be needed.

A question that cannot be answered from the relatively short DYAMOND simulations is whether the spread in the water-vapor–lapse-rate 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 feedback depends on the base-state RH as suggested by Bourdin et al. (2021), reducing the inter-model spread in present-day RH should also reduce the spread in the feedback. On the other hand, the water-vapor–lapse-rate feedback depends on how much RH changes under warming. Given that the present-day RH is better constrained in GSRMs, it seems unlikely that the spread in the RH response is increased. This is to be verified once model simulations at higher SSTs are available.

Although RH differences are reduced in the DYAMOND ensemble, they still cause a spread of $1.2\,\mathrm{W m^{-2}}$ 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 σ(RH) with radiative kernels (the sensitivity of clear-sky 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 σ(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 clear-sky 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 σ(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, which might be indicative of model differences in convective behavior. The large RH differences cause the inter-model OLR spread to maximize in this region, although the sensitivity of clear-sky OLR to RH perturbations is moderate.

3. The mid troposphere of dry regimes. In this region the model agreement in RH is remarkably good. The inter-model standard deviation σ(RH) is 1–3% RH and hence less than half of the standard deviation in moist regimes. However, the sen-

sitivity of clear-sky 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 spread in clear-sky OLR 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 σ(RH) exceeding 8% near the tropopause. However, the sensitivity of clear-sky OLR 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, mask some of the clear-sky effect (e.g. Soden et al., 2008) and hence reduce the radiative impact of the RH differences in the mid troposphere. This highlights even more the importance of the dry regimes, where high clouds are rare.

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 the processes controlling 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 σ(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 $\mathrm{OLR_c}$ anomaly in these regions is rather small (Figure 6). 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 $\mathrm{OLR_c}$ anomalies almost fully compensate in the tropical mean (Figure 6).

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

To obtain the radiative kernels $\mathbf{K^e}$ and $\mathbf{K^T}$ for a given block in moisture space, $\mathrm{OLR_c}$ 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 chosen perturbation sizes lie within the range for which the assumption of linearity around the base state is valid. Within this range the calculated kernels are independent of the exact perturbation size.

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 1) 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 it is likely that 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 1. Such SST anomalies have a stronger impact in the dry regions because the surface component of $\mathrm{OLR_c}$ is larger there than in moist regions. The largest deviations between true and approximated $\mathrm{OLR_c}$ 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, anomalies in $\mathrm{OLR_c}$ can also be approximated from RH anomalies and a RH kernel (Equation 1). There are two ways to define a RH kernel by varying either e or T (Equation 3), which we refer to as $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ and $\mathbf{K}^{\mathrm{RH},\mathbf{T}}$, respectively. Our main analysis is based on $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ because it approximates the anomalies in $\mathrm{OLR_c}$ more accurately. The largest deviations from true (directly calculated) $\mathrm{OLR_c}$ anomalies occur for SAM in the lowest IWV percentiles, for ARPEGE-NH in high percentiles and for ICON in all percentiles (Figure 7c). The inter-model standard deviation $\sigma(\mathrm{OLR})$ is well reproduced with the approximated $\mathrm{OLR_c}$ (Figure 7d), 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 $\mathrm{OLR_c}$ approximated from RH anomalies matches the true $\mathrm{OLR_c}$ much better than the one approximated from e and T anomalies.

For completeness Figure B2 shows $\mathbf{K}^{\mathrm{RH},\mathbf{T}}$ and the $\mathrm{OLR_c}$ anomalies approximated using this version of the RH kernel. $\mathbf{K}^{\mathrm{RH},\mathbf{T}}$ takes on larger absolute values than $\mathbf{K}^{\mathrm{RH},\mathbf{e}}$ (cf. Figure 7a, note the different colour scales in Figures 7 and B2), i.e. a 1% increase in RH causes a larger decrease in clear-sky 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 $\mathrm{OLR_c}$ 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 $\mathrm{OLR_c}$ 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.

The DYAMOND data and further management was provided by the German Climate Computing Center (DKRZ) and made available through the projects ESiWACE and ESiWACE2 (`https://www.esiwace.eu/services/dyamond`). 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 (`https://cera-www.dkrz.de/WDCC/ui/cerasearch/`).

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. We also thank Steven Sherwood and one anonymous reviewer for thoughtful and stimulating comments.

The authors declare no 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

Becker, T., & Wing, A. A. (2020, oct). Understanding the extreme spread in climate sensitivity within the radiative-convective equilibrium model intercomparison project. *Journal of Advances in Modeling Earth Systems*, *12*(10). doi: 10.1029/2020MS002165

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

Bony, S., Colman, R., Kattsov, V., Allan, R., Bretherton, C., Dufresne, J.-L., ... Webb, M. (2006, aug). How well do we understand and evaluate climate change feedback processes? *Journal of Climate*, *19*(15), 3445–3482. doi: 10.1175/JCLI3819.1

Bourdin, S., Kluft, L., & Stevens, B. (2021, apr). Dependence of climate sensitivity on the given distribution of relative humidity. *Geophysical Research Letters*. doi: https://doi.org/10.1029/2021GL092462

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

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

Clough, S. A., Iacono, M. J., & Moncet, J.-L. (1992). Line-by-line calculations of atmospheric fluxes and cooling rates: Application to water vapor. *Journal of Geophysical Research*, *97*(D14), 15761. doi: 10.1029/92JD01419

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

Derbyshire, S., Beau, I., Bechtold, P., Grandpeix, J.-Y., Piriou, J.-M., Redelsperger, J.-L., & Soares, P. (2004, oct). Sensitivity of moist convection to environmental humidity. *Quarterly Journal of the Royal Meteorological Society*, *130*(604), 3055–3079. doi: 10.1256/qj.03.130

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

Hartmann, D., & Larson, K. (2002, oct). An important constraint on tropical cloud - climate feedback. *Geophysical Research Letters*, *29*(20), 12–1–12–4. doi: 10.1029/2002GL015835

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

Jeevanjee, N., & Fueglistaler, S. (2020, jan). Simple spectral models for atmospheric radiative cooling. *Journal of the Atmospheric Sciences*, *77*(2), 479–497. doi: 10.1175/JAS-D-18-0347.1

Jensen, E. J., Read, W. G., Mergenthaler, J., Sandor, B. J., Pfister, L., & Tabazadeh, A. (1999, aug). High humidities and subvisible cirrus near the tropical tropopause. *Geophysical Research Letters*, *26*(15), 2347–2350. doi: 10.1029/1999GL900266

John, V., Buehler, S., von Engeln, A., Eriksson, P., Kuhn, T., Brocard, E., & Koenig-Langlo, G. (2006, oct). Understanding the variability of clear-sky outgoing long-wave radiation based on ship-based temperature and water vapour measurements. *Quarterly Journal of the Royal Meteorological Society*, *132*(621), 2675–2691. doi: 10.1256/qj.05.70

John, V., & Soden, B. (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

Koll, D. D. B., & Cronin, T. W. (2018, sep). Earth's outgoing longwave radiation linear due to h2o greenhouse effect. *Proceedings of the National Academy of Sciences*, *115*(41), 10293–10298. doi: 10.1073/pnas.1809868115

Luo, Z., & Rossow, W. (2004). Characterizing tropical cirrus life cycle, evolution, and interaction with upper-tropospheric water vapor using lagrangian trajectory analysisof satellite observations. *Journal of Climate*, *17*(23), 4541–4563. doi: 10.1175/3222.1

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

Miura, H., Satoh, M., Nasuno, T., Noda, A. T., & Oouchi, K. (2007, dec). A madden-julian oscillation event realistically simulated by a global cloud-resolving model. *Science*, *318*(5857), 1763–1765. doi: 10.1126/science.1148443

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

Petty, G. (2006). *A first course in atmospheric radiation* (2nd ed. ed.). Sundog Publishing Madison, Wisconsin.

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

Pincus, R., Mlawer, E. J., Oreopoulos, L., Ackerman, A. S., Baek, S., Brath, M., ... Schwarzkopf, D. M. (2015, jul). Radiative flux and forcing parameterization error in aerosol-free clear skies. *Geophysical Research Letters*, *42*(13), 5485–5492. doi: 10.1002/2015GL064291

Po-Chedley, S., Armour, K. C., Bitz, C. M., Zelinka, M. D., Santer, B. D., & Fu, Q. (2018, mar). Sources of intermodel spread in the lapse rate and water vapor feedbacks. *Journal of Climate*, *31*(8), 3187–3206. doi: 10.1175/JCLI-D-17-0674.1

Po-Chedley, S., Zelinka, M., Jeevanjee, N., Thorsen, T., & Santer, B. (2019). Climatology explains intermodel spread in tropical upper tropospheric cloud and relative humidity response to greenhouse warming. *Geophysical Research Letters*, *46*(22), 13399–13409. doi: 10.1029/2019GL084786

Retsch, M. H., Mauritsen, T., & Hohenegger, C. (2019, jul). Climate change feedbacks in aquaplanet experiments with explicit and parametrized convection for horizontal resolutions of 2,525 up to 5 km. *Journal of Advances in Modeling Earth Systems*, *11*(7), 2070–2088. doi: 10.1029/2019MS001677

Romps, D. (2014, sep). An analytical model for tropical relative humidity. *Journal of Climate*, *27*(19), 7432–7449. doi: 10.1175/JCLI-D-14-00255.1

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., Ingram, W., Tsushima, Y., Satoh, M., Roberts, M., Vidale, P., & O'Gorman, P. (2010). Relative humidity changes in a warmer climate. *Journal of Geophysical Research*, *115*(D9), D09104. doi: 10.1029/2009JD012585

Shine, K. P., & Sinha, A. (1991, dec). Sensitivity of the earth's climate to height-dependent changes in the water vapour mixing ratio. *Nature*, *354*(6352), 382–384. doi: 10.1038/354382a0

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

Soden, B., & Held, I. (2006, jul). An assessment of climate feedbacks in coupled ocean–atmosphere models. *Journal of Climate*, *19*(14), 3354–3360. doi: 10.1175/JCLI3799.1

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

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., 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

Tomita, H., Miura, H., Iga, S., Nasuno, T., & Satoh, M. (2005). A global cloud-resolving simulation: Preliminary results from an aqua planet experiment. *Geophysical Research Letters*, *32*(8). doi: 10.1029/2005GL022459

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

Vömel, H., Oltmans, S. J., Johnson, B. J., Hasebe, F., Shiotani, M., Fujiwara, M., ... Enriquez, H. (2002). Balloon-borne observations of water vapor and ozone in the tropical upper troposphere and lower stratosphere. *Journal of Geophysical Research*, *107*(D14). doi: 10.1029/2001JD000707

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

Wing, A. A., Stauffer, C. L., Becker, T., Reed, K. A., Ahn, M.-S., Arnold, N. P., ... Zhao, M. (2020). Clouds and convective self-aggregation in a multimodel ensemble of radiative-convective equilibrium simulations. *Journal of Advances in Modeling Earth Systems*, *12*(9). doi: 10.1029/2020MS002138

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

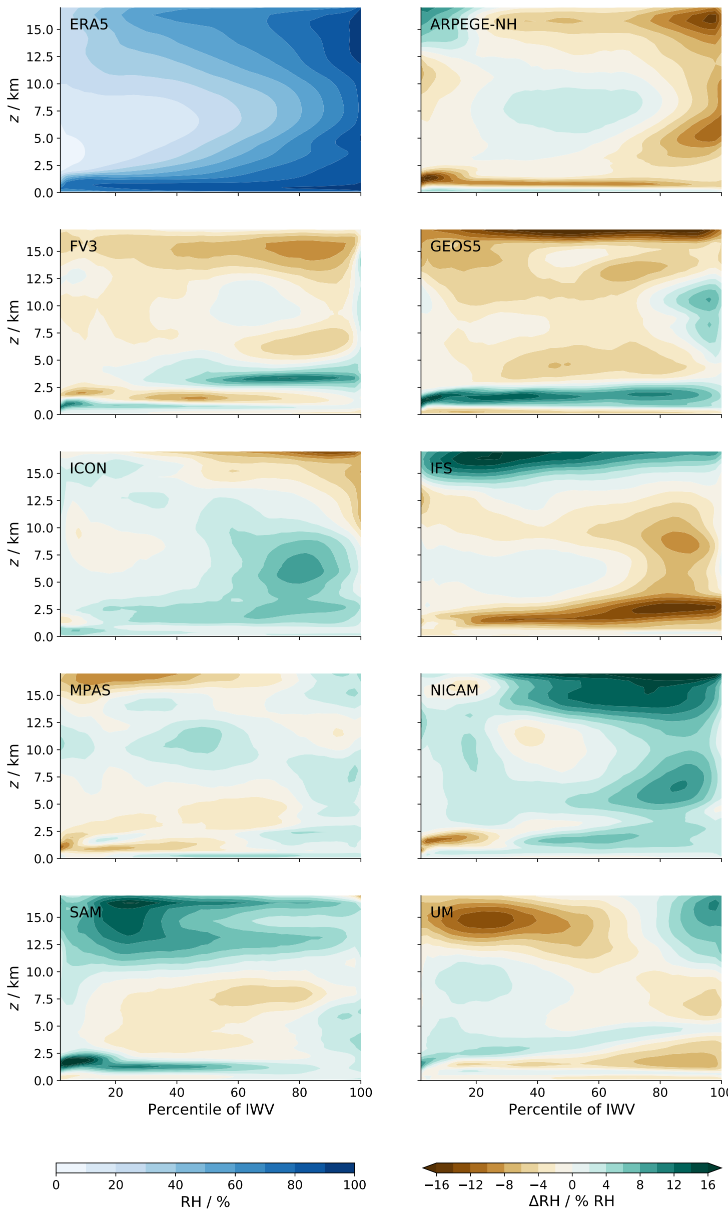

Figure A1. RH anomalies of DYAMOND models in moisture space. The upper left panel shows the ERA5 RH distribution in moisture space, remaining panels show the deviation from the ERA5 RH for each model.

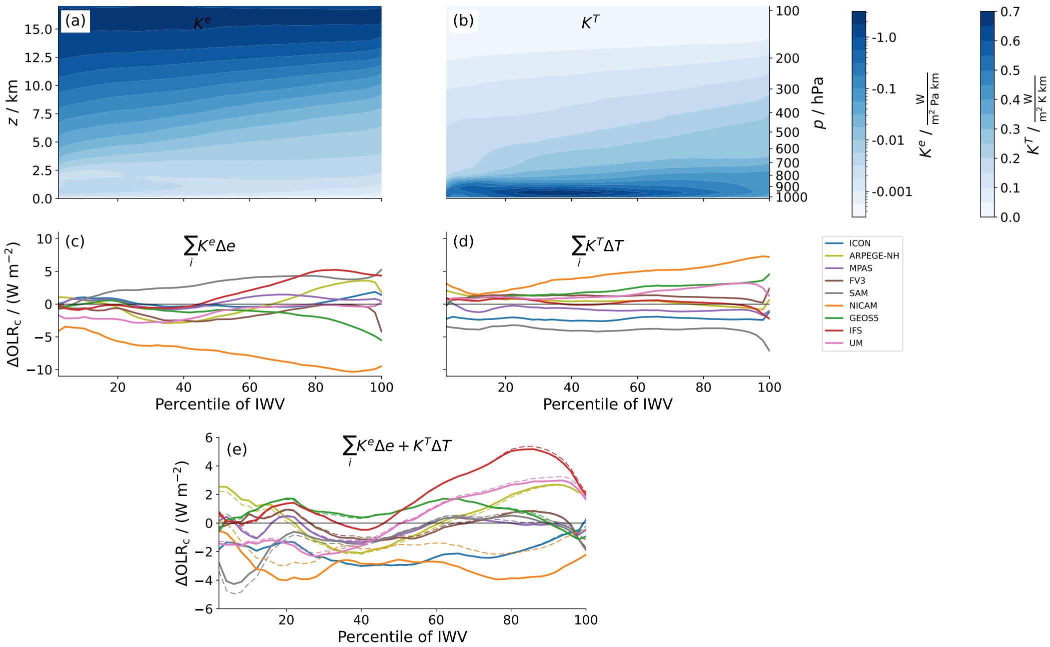

Figure B1. Clear-sky OLR anomalies in the DYAMOND models approximated with the kernel method. (a) Water vapor response kernel $\mathbf{K^e}$ showing the sensitivity of clear-sky OLR to a change of 1 Pa in water vapor pressure e in a 1 km layer. Note the logarithmic colour scale. (b) Temperature response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

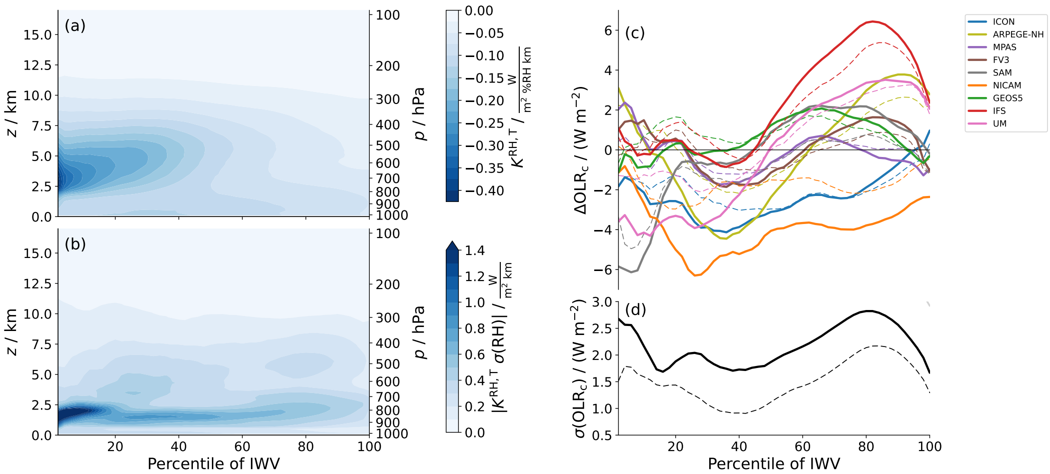

Figure B2. As Figure 7 but based on $\mathbf{K}^{\text{RH},\mathbf{T}}$. Note that the colour scale in (a) and (b) is different from Figure 7 since $\mathbf{K}^{\text{RH},\mathbf{T}}$ takes on more negative values than $\mathbf{K}^{\text{RH},\mathbf{e}}$.