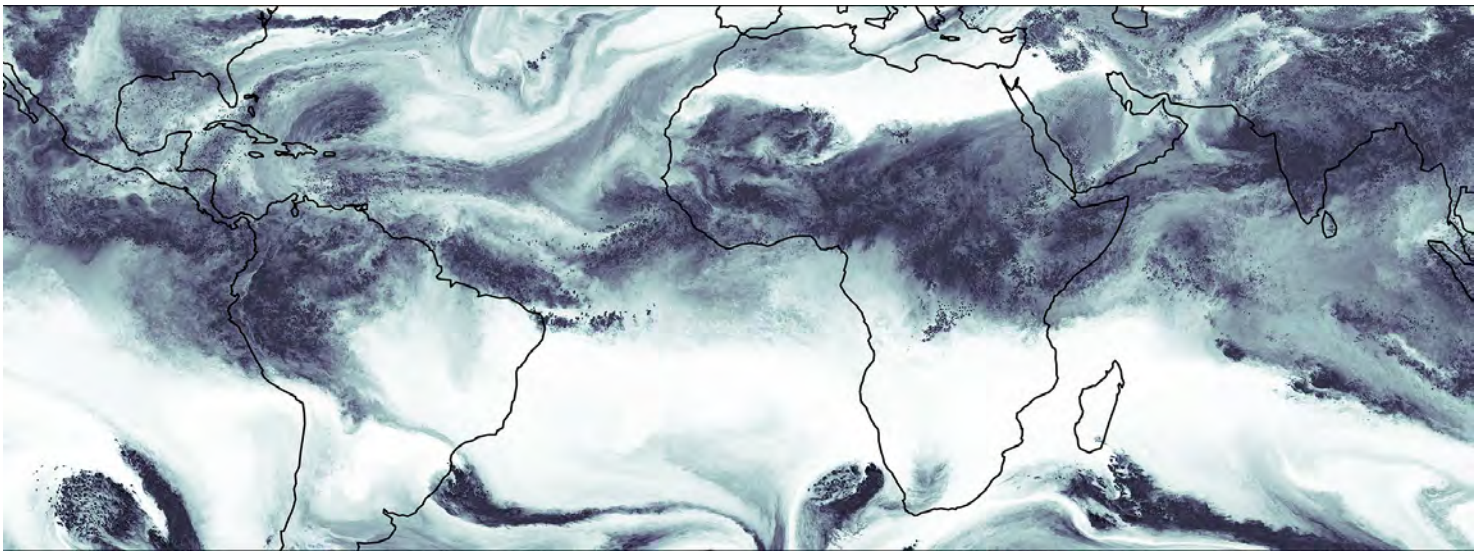
MAX-PLANCK-INSTITUT
FÜR METEOROLOGIE

On the Uncertainty in Modelling Tropical Relative Humidity



Theresa Claudia Lang

Hamburg 2023

Berichte zur Erdsystemforschung
Reports on Earth System Science

263
2023

Hinweis

Die Berichte zur Erdsystemforschung werden vom Max-Planck-Institut für Meteorologie in Hamburg in unregelmäßiger Abfolge herausgegeben.

Sie enthalten wissenschaftliche und technische Beiträge, inklusive Dissertationen.

Die Beiträge geben nicht notwendigerweise die Auffassung des Instituts wieder.

Die "Berichte zur Erdsystemforschung" führen die vorherigen Reihen "Reports" und "Examensarbeiten" weiter.

Anschrift / Address

Max-Planck-Institut für Meteorologie
Bundesstrasse 53
20146 Hamburg
Deutschland

Tel./Phone: +49 (0)40 4 11 73 - 0
Fax: +49 (0)40 4 11 73 - 298

name.surname@mpimet.mpg.de
www.mpimet.mpg.de

Notice

The Reports on Earth System Science are published by the Max Planck Institute for Meteorology in Hamburg. They appear in irregular intervals.

They contain scientific and technical contributions, including PhD theses.

The Reports do not necessarily reflect the opinion of the Institute.

The "Reports on Earth System Science" continue the former "Reports" and "Examensarbeiten" of the Max Planck Institute.

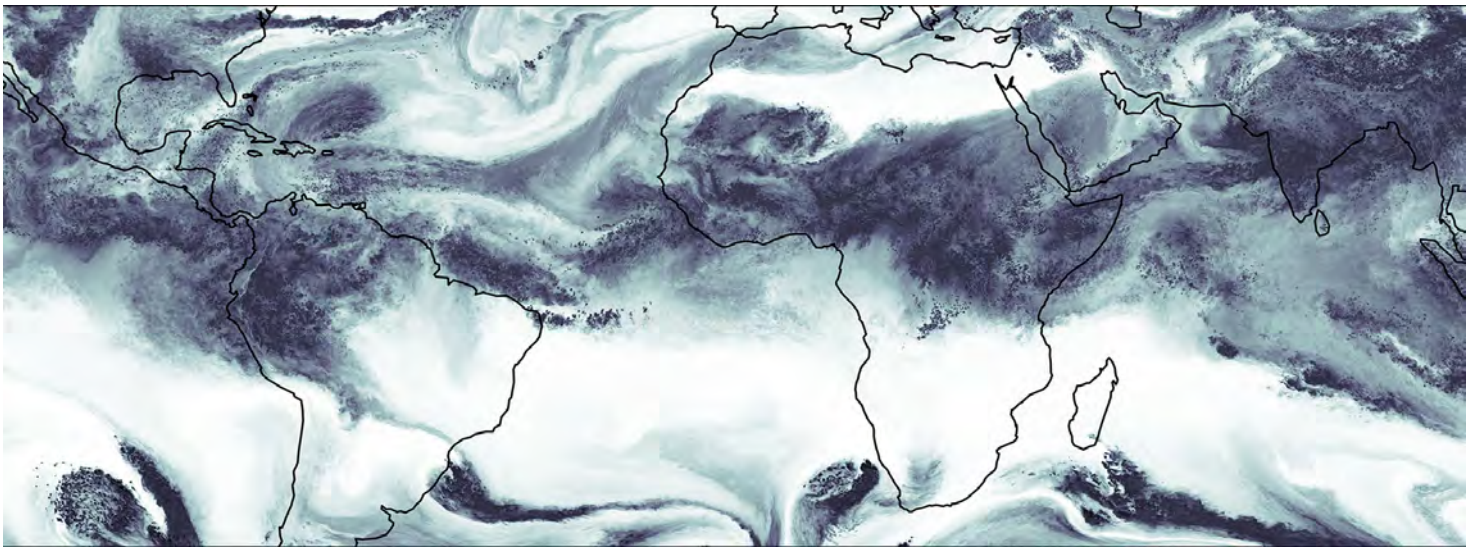
Layout

Bettina Diallo and Norbert P. Noreiks
Communication

Copyright

Photos on the back from left to right:
Christian Klepp, Jochem Marotzke,
Christian Klepp, Clotilde Dubois,
Christian Klepp, Katsumasa Tanaka

On the Uncertainty in Modelling Tropical Relative Humidity



Theresa Claudia Lang

Hamburg 2023

Theresa Claudia Lang
aus Burgthann, Deutschland

Max-Planck-Institut für Meteorologie
The International Max Planck Research School on Earth System Modelling
(IMPRS-ESM)
Bundesstrasse 53
20146 Hamburg

Universität Hamburg
Erdsystemwissenschaften
Bundesstr. 55
20146 Hamburg

Tag der Disputation: 24. Februar 2023

Folgende Gutachter empfehlen die Annahme der Dissertation:
Dr. Ann Kristin Naumann
Prof. Dr. Stefan A. Buehler

Vorsitzender des Promotionsausschusses:
Prof. Dr. Hermann Held

Dekan der MIN-Fakultät:
Prof. Dr.-Ing. Norbert Ritter

Titelgrafik von Theresa Lang: *Mid-tropospheric relative humidity in the tropics simulated by the storm-resolving model ICON*

Berichte zur Erdsystemforschung / Max-Planck-Institut für Meteorologie 263
Reports on Earth System Science / Max Planck Institute for Meteorology 2023

ISSN 1614-1199 - doi: 10.17617/2.3501977

ABSTRACT

The overarching goal of this dissertation is to advance the understanding of uncertainties in modelling the distribution of tropical relative humidity. Climate models struggle to consistently simulate this distribution and its change with warming, which introduces uncertainty in the present-day radiation budget as well as the projected future warming. The sources of inter-model differences in the humidity distribution are, however, poorly understood. In conventional climate models, many processes involved in establishing the humidity distribution are not resolved, but instead crudely represented by parameterizations. These processes include vertical mixing by convection, small-scale turbulent mixing and cloud microphysical processes. As they are all coupled through the convective parameterization, it is difficult to relate humidity differences across models to physical processes. This changes with global storm-resolving models (GSRMs), the latest generation of global models. They are run at much higher horizontal resolution than conventional models, which allows to explicitly simulate deep convection rather than parameterizing it.

Whether resolving deep convection leads to a reduction of humidity differences across models is not yet known. In my first study, I investigate this based on a first 40-day comparison of nine atmospheric GSRMs. I find that, in the free troposphere, the model spread in tropical mean relative humidity is approximately halved compared to conventional atmospheric models, indicating that explicitly simulating deep convection leads to a more consistent representation of relative humidity. However, I also show that the remaining differences translate into a significant uncertainty of the clear-sky radiation budget. By examining humidity differences in different tropical moisture regimes, I find that comparably modest differences in dry regimes have a similar clear-sky radiative impact as larger differences in moist regimes. Although humidity differences are largest in the upper troposphere, I identify the tropical lower and mid free troposphere as the altitude region where a reduction of differences would be most beneficial from a radiative perspective. Particularly in these altitudes, a better understanding of the processes controlling the humidity differences is needed.

In my second study, I therefore aim to narrow down the sources of uncertainties in modelling the distribution of tropical relative humidity with GSRMs. In seven 45-day experiments I examine how much the relative humidity simulated by a GSRM is affected by changes in model resolution and in the paramterizations of microphysics and turbulence. To investigate the physical mechanisms behind the changes in humidity, I perform trajectory-based reconstructions of the relative humidity distribution. As a null-hypothesis I use the last-saturation paradigm, according to which the water vapor content of an air parcel only depends on the temperature it had when it was last saturated. I find that mid-tropospheric relative humidity is more sensitive to changes in parameterizations than in model resolution, suggesting that model physics represent a major source of humidity

differences across GSRMs. The relative humidity in moist tropical regions is disproportionately sensitive to the strength of vertical turbulent mixing in the inner tropics, which impacts relative humidity through its effect on last-saturation temperature rather than its effect on the evolution of humidity since last-saturation. The humidity of the dry regions depends on the pathways of exchange with the extra-tropics. More research is needed to understand these pathways, and the uncertainties in modelling them, in order to build trust in the simulated humidity and its change with warming in the radiatively important dry regions.

ZUSAMMENFASSUNG

Das übergeordnete Ziel dieser Dissertation ist es, ein besseres Verständnis der Unsicherheiten bei der Modellierung der Verteilung der relativen Feuchte in den Tropen zu erlangen. Klimamodelle stellen diese Verteilung und ihre Veränderung im Zuge einer Erwärmung nicht konsistent dar, was zu Unsicherheiten im Strahlungsbudget sowie in der prognostizierten Erwärmung führt. Die Ursachen für Modellunterschiede in der Feuchteverteilung sind jedoch nicht ausreichend verstanden. In konventionellen Modellen sind viele, für die Feuchteverteilung relevante, Prozesse nicht aufgelöst, sondern werden nur vereinfacht in Form von Parametrisierungen dargestellt. Dazu gehören vertikale Mischung durch Konvektion, kleinskalige turbulente Mischung sowie mikrophysikalische Prozesse. Da sie alle durch die Konvektionsparametrisierung gekoppelt sind, ist es schwierig, die Modellunterschiede in der Feuchte mit physikalischen Prozessen in Verbindung zu setzen. Dies ändert sich mit globalen sturmauflösenden Modellen (GSRMs), der neuesten Generation globaler Modelle. Diese verfügen über eine viel höhere horizontale Auflösung als herkömmliche Modelle und sind dadurch in der Lage hohe Konvektion explizit zu simulieren, statt sie zu parametrisieren.

Ob das Auflösen hoher Konvektion zu einer Verringerung der Feuchteunterschiede zwischen Modellen führt, ist noch nicht bekannt. In meiner ersten Studie untersuche ich dies basierend auf einem ersten 40-tägigen Vergleich von neun atmosphärischen GSRMs. Ich zeige, dass die Modellstreuung in der gemittelten tropischen relativen Feuchte im Vergleich zu herkömmlichen Klimamodellen in der freien Troposphäre etwa halbiert ist. Dies deutet darauf hin, dass das Auflösen tiefer Konvektion zu einer konsistenteren Darstellung der relativen Feuchte führt. Ich zeige jedoch auch, dass die verbleibenden Feuchteunterschiede zu einer erheblichen Unsicherheit des Strahlungsbudges bei wolkenfreier Atmosphäre führen. Eine Untersuchung der Modellunterschiede in verschiedenen Feuchte-Regimen der Tropen zeigt, dass vergleichsweise kleine Feuchteunterschiede in trockenen Regimen bei wolkenfreier Atmosphäre eine ähnliche Strahlungswirkung haben wie größere Unterschiede in feuchten Regimen. Außerdem zeige ich, dass, obwohl Feuchtesunterschiede in der oberen Troposphäre am größten sind, eine Verringerung der Unterschiede in der unteren und mittleren freien Troposphäre aus Strahlungssicht am wirksamsten wäre. Insbesondere in diesen Höhen ist deshalb ein besseres Verständnis derjenigen Prozesse erforderlich, die die Feuchteunterschiede verursachen.

In meiner zweiten Studie versuche ich deshalb, die Unsicherheitsquellen bei der Modellierung der tropischen Feuchteverteilung mit GSRMs einzugrenzen. In sieben 45-tägigen Experimenten untersuche ich, wie stark die von einem GSRM simulierte relative Feuchte durch Änderungen in der Modellformulierung beeinflusst wird. Diese Änderungen betreffen sowohl die Modellauflösung als auch die Parametrisierungen von Mikrophysik und Turbulenz, die auf sturmauflösenden Skalen bestehen bleiben. Um zu untersuchen welche physikalischen Mechanismen

für die Feuchteänderungen verantwortlich sind, rekonstruiere ich die Feuchteverteilung basierend auf Rückwärts-Trajektorien. Als Nullhypothese nutze ich dabei das Modell der letzten Sättigung, demzufolge der Wasserdampfgehalt eines Luftpakets nur von der Temperatur bei der letzten Sättigung abhängt. Ich zeige, dass die relative Luftfeuchtigkeit in der mittleren Troposphäre stärker auf Änderungen der Parametrisierungen als auf Änderungen der Modellauflösung reagiert. Dies deutet darauf hin, dass die Modellphysik eine Hauptquelle für die Feuchteunterschiede zwischen GSRMs darstellt. In feuchten tropischen Regionen ist die relative Feuchte besonders empfindlich gegenüber der Stärke der vertikalen turbulenten Mischung in den inneren Tropen. Diese beeinflusst die relative Feuchte durch Änderungen in der Sättigungstemperatur, nicht durch Änderungen der Feuchteentwicklung nach der letzten Sättigung. In trockenen Regionen hängt die relative Feuchte von den Austauschpfaden mit den Außertropen ab. Weitere Forschung ist nötig um diese Austauschpfade, und die Unsicherheiten bei deren Modellierung, zu verstehen, und so die Belastbarkeit der simulierten Feuchte und deren Änderung bei Erwärmung zu erhöhen.

PUBLICATIONS RELATED TO THIS DISSERTATION

The following two first-author publications are part of this dissertation and included in the appendix:

APPENDIX A

Lang, T., Naumann, A. K., Stevens, B., and Buehler, S. A. (2021). Tropical free-tropospheric humidity differences and their effect on the clear-sky radiation budget in global storm-resolving models. *Journal of Advances in Modeling Earth Systems*, 13, e2021MS002514. https://doi.org/10.1029/2021MS002514

APPENDIX B

Lang, T., Naumann, A. K., Buehler, S. A., Stevens, B., Schmidt, H. and Aemisegger, F. (2022). How model uncertainties influence tropical humidity in global storm-resolving simulations. In review for *Journal of Advances in Modeling Earth Systems*.

As a PhD student I contributed to four further publications. The first one is a first-author paper that contains the results of my master thesis and some further analysis, to the other three I contributed as a co-author. More information about these publications and my specific contributions can be found in Chapter 3.

Lang, T., Buehler, S. A., Burgdorf, M., Hans, I. and John, V. O. (2020). A new climate data record of upper-tropospheric humidity from microwave observations. *Scientific Data*, 7, 218. https://doi.org/10.1038/s41597-020-0560-1.

Stevens, B., Bony, S., Farrell, D., and 290 co-authors (incl. **Lang, T.**) (2021). EUREC4A, *Earth System Science Data*, 13, 4067–4119. https://doi.org/10.5194/essd-13-4067-2021

Konow, H., Ewald, F., George, G., and 35 co-authors (incl. **Lang, T.**) (2021). EUREC4A's HALO, *Earth System Science Data*, 13, 5545–5563. https://doi.org/10.5194/essd-13-5545-2021

Shi, L., Schreck III, C. J., John, V. O., Chung, E.-S., **Lang, T.**, Buehler, S. A., and Soden, B. J. Assessing the consistency of satellite derived upper tropospheric humidity measurements, Accepted for *Atmospheric Measurement Techniques*. https://doi.org/10.5194/amt-2022-204

ACKNOWLEDGEMENTS

During my PhD I received a lot of help and support from many different people, who I want to thank here.

First and foremost, I would like to thank my supervisors Ann Kristin Naumann and Stefan Bühler for giving me the opportunity to conduct this research, for taking time and interest in my work, and for giving me a lot of freedom in developing my ideas. Thank you for giving me the possibility to present my research at several conferences, and particularly also to participate in the EUREC^{4}A field campaign, which was an unforgettable experience. I am also very grateful to Bjorn Stevens, who provided me with great ideas and thought-provoking impulses during many discussions. Furthermore, I want to thank Hauke Schmidt for welcoming me in his group and co-supervising me towards the end of my PhD, as well as Johanna Baehr for guiding us through the panel meetings so efficiently.

During my PhD I had the privilege to be part of two groups at UHH and MPI: the *Drivers of Tropical Circulation* group and the *Climate, Radiation, Remote Sensing* group. I thank all members of both groups for their constructive feedback and for asking questions that helped me to sharpen my research. I am very grateful to have met Jule Radtke, with whom I shared my entire PhD journey, as well as the life between two institutes. Thank you for all your scientific and moral support during this time. I am also very thankful to my colleague Lukas Kluft, for always taking time to answer any questions related to programming or radiation, but especially for spending so many hours on helping me out with my simulations. Thanks to your support I did not only get ICON to run, but also had a lot of fun.

Moreover, I would like to thank my wonderful colleagues and office mates, especially Jule, Marc, Simon, Lukas and Leo, for countless discussions about work and life, for refreshing lunch breaks, for many evenings at Planten un Blomen and the Christmas market, and a lot of other fun activities, even virtually in times of Covid. I also thank you for your thoughtful comments on this thesis.

Last but not least, I want to thank my family and friends. Thank you for repeatedly reminding me that work is not everything, but especially for always supporting me, although I am so far away from most of you. It remains to thank Tobi. Thank you for being so incredibly patient and understanding, especially in stressful times, for cheering me up, for celebrating every tiny achievement with me, and for just always being there. Because of you, I look into the future full of confidence and joy.

CONTENTS

Part I

UNIFYING ESSAY

1

MOTIVATION: REDUCING THE UNCERTAINTY IN MODELLING TROPICAL HUMIDITY

Earth's climate is strongly shaped by water vapor in the atmosphere. Besides its influence on atmospheric circulations through the release of latent heat, water vapor plays a key role in establishing the radiative balance of the planet – both directly and indirectly through the formation of clouds. Climate models, however, struggle to consistently simulate the atmosphere's water vapor distribution. In this thesis, I investigate the physical mechanisms underlying the model spread in humidity and the uncertainty this spread introduces in the clear-sky radiation budget.

Water vapor is the most important greenhouse gas in our atmosphere. Like all greenhouse gases it interacts with the longwave thermal radiation emitted from the Earth's surface. As a consequence, in those spectral regions of the emission spectrum that are dominated by water vapor absorption, the radiation emitted to space originates from within the troposphere rather than the surface. Since the troposphere generally has a lower temperature than the surface, this effectively reduces the total outgoing longwave radiation (OLR).

OLR is particularly sensitive to changes in atmospheric relative humidity (e.g. Ingram, 2010), which denotes the ratio of the actual water vapor content to its thermodynamic equilibrium value. The latter is an increasing function of temperature. In the water-vapor-dominated parts of the spectrum, the absorption by water vapor determines the altitude layer from which radiation is emitted to space. The temperature of this layer sets the thermal emission. An increase in atmospheric temperature increases the emission and hence OLR. However, if the temperature increase is accompanied by an increase in water vapor concentration, the emission layer shifts to higher levels of lower temperature, which in turn reduces OLR. If the atmospheric temperature profile and water vapor concentration jointly increase or decrease such that the relative humidity stays unchanged, their competing effects on OLR largely compensate. Therefore, the relative humidity is a valuable proxy that measures the combined effect of temperature and water vapor changes on OLR.

Relative humidity does not only influence the instantaneous clear-sky radiation budget, but also the clear-sky climate feedback. The feedback determines the amount of global surface warming needed to restore radiative equilibrium after a radiative forcing such as an increase in greenhouse gas concentrations. There is a wide agreement that, as the Earth's surface warms, the vertical thermal structure of the atmosphere (in particular the lapse-rate, i.e. the temperature decrease with height) and the water vapor concentration of the atmosphere will

respond in such a way that relative humidity stays roughly unchanged (e.g. Colman and Soden, 2021; Held and Soden, 2000). Compared to a basic Planck response, which describes a vertically uniform warming at constant absolute humidity, the non-uniform warming at constant relative humidity reduces the efficiency at which the Earth can cool to space and therefore requires a stronger surface warming to restore radiative equilibrium. This represents a positive feedback, which is often expressed as the combined water-vapor–lapse-rate feedback. Together with the Planck feedback the water-vapor–lapse-rate feedback describes the Earth's clear-sky response to a radiative forcing to a large extent.

Although there is agreement that the water-vapor–lapse-rate feedback is positive, its exact strength depends sensitively on the distribution of free-tropospheric relative humidity and its response to warming. On the one hand, even small changes of relative humidity with warming can significantly alter the clear-sky feedback by changing the partial cancellation between water-vapor and lapse-rate feedbacks (Bony et al., 2006). Such changes in relative humidity are expected to occur particularly in the tropics (O'Gorman and Muller, 2010; Vial et al., 2013). On the other hand, recent studies using 1D radiative-convective equilibrium models show that the tropical clear-sky feedback is also influenced by the present-day relative humidity (Bourdin et al., 2021; Kluft et al., 2019; McKim et al., 2021). The mechanism behind this is a narrowing of the water vapor window, which denotes the region of the emission spectrum in which radiation emitted from the Earth's surface can escape to space because absorption by water vapor is weak. This narrowing occurs at high surface temperatures characteristic for the tropics and makes the feedback dependent on the present-day relative humidity, even if it is assumed to stay unchanged under warming. Due to the mentioned implications for the clear-sky feedback, it is crucial that climate models correctly capture atmospheric relative humidity and its evolution under warming for them to reliably project future warming.

The models that are currently most widely used to study the temperature response to a given radiative forcing are General Circulation Models (GCMs; IPCC, 2021). These models depict the atmosphere using a three-dimensional grid with a horizontal grid spacing that is typically on the order of 100 km. However, many processes relevant for climate act on scales smaller than a GCM grid cell. These processes are parameterized. That is, they are represented in a simplified manner, by relating their temporally and spatially averaged effect to resolved (large-scale) variables using semi-empirical relations. An important parameterized process in GCMs is moist convection, which accounts for most of the vertical transport of heat, momentum, and moisture in the tropics. Other parameterized processes include turbulent mixing and cloud microphysics. Parameterizations, in particular the parameterization of deep convective clouds, are known to represent a major uncertainty in GCMs (Stevens and Bony, 2013a) and are commonly used to tune the models to match the observed state of the atmosphere (Mauritsen et al., 2012).

All of the parameterized processes mentioned above are potentially important for the tropospheric humidity distribution, which is determined by advection of water vapor on all scales as well as cloud microphysical processes, such as evaporation of cloud particles (liquid or ice) and precipitation (e.g. Emanuel and Pierrehumbert, 1996). Despite the uncertainties associated with these parameterizations, GCMs to a first approximation capture the large-scale humidity distribution in the tropics because it is mainly driven by the large-scale circulation, which GCMs are able to resolve (e.g. Dessler and Sherwood, 2000; Sherwood, 1996). Nevertheless, there is a considerable spread in relative humidity as well as in its change with warming across GCMs (e.g. Po-Chedley et al., 2019). These differences in the relative humidity response to warming control the current spread in model-based estimates of the tropical water-vapor–lapse-rate feedback, which accounts for a large part of the spread in the global clear-sky feedback (Po-Chedley et al., 2018; Vial et al., 2013). Therefore, it is important to reduce the uncertainty in tropical relative humidity and its change with warming. However, a detailed understanding of the underlying physical mechanisms is difficult in GCMs due to the number of parameterized processes, and due to the fact that they are all coupled through the convective parameterization.

One approach to reduce the uncertainty associated with parameterizations, is the development of global storm-resolving models (GSRMs), also referred to as global cloud-resolving or convection-permitting models (Satoh et al., 2019). These models are run at horizontal grid spacings on the order of a few kilometers, and are therefore capable of explicitly resolving the vertical transport associated with deep convection. Microphysical processes and turbulent mixing still need to be parameterized in GSRMs. However, unlike in models with a convective parameterization, they are now directly linked to the resolved circulation. Therefore, these models promise new insights into the importance of these processes, particularly regarding the tropical humidity distribution as investigated in this work.

GSRMs are still in an early phase of development. Running global simulations on the kilometer-scale comes at a high computational cost and has therefore only become feasible in recent years. Global storm-resolving simulations are therefore currently limited to rather short time periods of weeks to months. Since the first GSRM was developed about 18 years ago (Tomita and Satoh, 2004), several other modeling groups have followed. This enabled a first, 40-day intercomparison of GSRMs in 2019, an initiative called DYnamics of the Atmospheric general circulation Modeled On Nonhydrostatic Domains (DYAMOND; Stevens et al., 2019).

In the first study encompassed in this thesis I make use of the DYAMOND multi-model ensemble to assess for the first time whether the spread in tropical relative humidity is reduced across GSRMs compared to GCMs. I quantify how much relative humidity differs within the DYAMOND ensemble and assess the relevance of these differences by calculating their effect on the clear-sky radiation budget using a radiative transfer model. I show that the inter-model spread in

tropical relative humidity is indeed reduced in GSRMs compared to GCMs, but that it still translates into a non-negligible spread in tropical mean clear-sky OLR.

To enable a further reduction of the humidity spread across GSRMs, its origins must be understood. In my second study, I therefore investigate the sources of uncertainties in modelling the distribution of tropical relative humidity. To this end, I examine how much and through which physical mechanisms the relative humidity in a GSRM is affected by changes in model resolution and the parameterizations of microphysics and turbulence. To understand the physical mechanisms, I perform trajectory-based reconstructions of relative humidity based on the so-called last-saturation paradigm, which claims that the water vapor content of an air parcel only depends on the temperature it had when it was last saturated. My results suggest that differences in parameterizations, rather than in model resolution, drive the humidity spread across GSRMs, mainly by changing the statistics of last-saturation.

This essay is structured as follows. Chapter 2 provides more background on how relative humidity affects clear-sky OLR and the climate feedback, as well as on the last-saturation paradigm. Chapter 3 briefly introduces several accompanying projects dealing with observations of humidity and clouds in the tropics, which I worked on within my PhD term. In Chapters 4 and 5 I describe the key methods and results of the two studies I conducted as part of my dissertation, which are summarized and followed up by my conclusions in Chapter 6.

2

BACKGROUND: THE TROPICAL HUMIDITY DISTRIBUTION AND ITS IMPORTANCE FOR RADIATION

This chapter provides background on the distribution of humidity in the tropical atmosphere as well as its importance for the clear-sky radiation budget and climate feedback. In my first study, I use the radiative impact as a measure of how relevant the identified humidity differences across GSRMs are. Furthermore, this chapter introduces the last-saturation paradigm, a simple theory of what determines the tropical relative humidity distribution. In my second study, I use this paradigm as a tool to understand the humidity differences between storm-resolving simulations that I assessed in my first study.

2.1 BASIC PROPERTIES OF THE ATMOSPHERIC HUMIDITY DISTRIBUTION

Despite its profound climatic impact, water accounts for only 2.5 % of the total mass of the atmosphere. Almost all of this water (about 99.5 %) is present in the form of vapor (Stevens and Bony, 2013b). In the presence of liquid water, such as cloud droplets, the partial pressure of water vapor in a given air parcel represents a balance between evaporation of liquid and condensation of vapor (the same is true for the presence of ice, where the balance is between sublimation and resublimation). When the two rates are equal, the water vapor pressure of the parcel is said to be saturated (although "equilibrated" would be a more appropriate term). The equilibrium or saturation water vapor pressure e^* only depends on temperature T, and is governed by the Clausius-Clapeyron differential equation:

$$\frac{\mathrm{d}e^*}{\mathrm{d}T} = \frac{L\, e^*}{R_v T^2} \quad , \tag{1}$$

where L is the latent heat of vaporization (or sublimation) and R_v is the gas constant for water vapor. As soon as the actual water vapor pressure e of the parcel reaches the equilibrium vapor pressure e^*, any further addition of water vapor leads to condensation, thereby bringing the vapor pressure back to equilibrium. The condensed water either remains suspended in the air as cloud droplets or aggregates and falls out as precipitation.

Equation 1 implies a roughly exponential increase of e^* with T. At 300 K (a typical temperature of the tropical surface, Figure 1a) e^* is 3664 Pa. At the freezing point e^* (over water) decreases to 614 Pa, whereas at 200 K (a typical temperature of the tropical tropopause) e^* (over ice) is only 0.3 Pa. This strong thermodynamic constraint on atmospheric humidity is reflected in tropical humidity profiles simulated by a GSRM (Figure 1b). Both tropical mean specific humidity q and its saturation value q^* decrease roughly exponentially by four orders of magnitude between the surface and the tropopause. Specific humidity denotes the mass of

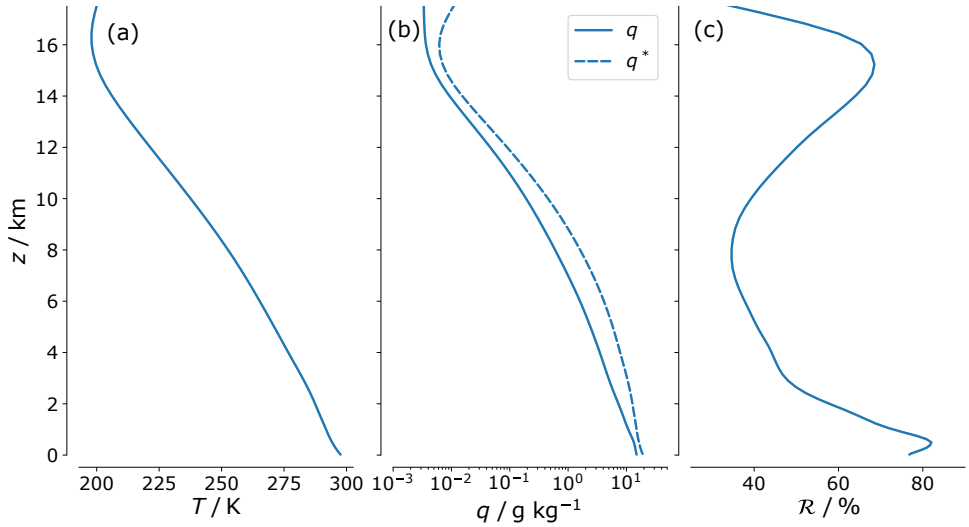


Figure 1: Tropical mean vertical profiles of (a) temperature T, (b) specific humidity q and saturation specific humidity q^*, and (c) relative humidity $\mathcal{R}$, averaged over a 45-day (27 June to 11 August) storm-resolving simulation with the ICOsahedral Nonhydrostatic model (ICON).

water vapor per unit mass of moist air and relates to the water vapor pressure e as

$$q = \frac{\frac{M_w}{M_d} e}{p - (1 - \frac{M_w}{M_d}) e} \approx \frac{M_w}{M_d} \frac{e}{p}, \tag{2}$$

where p is the partial pressure of dry air and M_w and M_d are the molar masses of water and dry air, respectively.

The Clausius-Clapeyron relation only provides a maximum value for the atmospheric water vapor concentration. Locally, the atmosphere can be highly subsaturated. The degree of subsaturation is expressed by the relative humidity $\mathcal{R}$, which is defined as the ratio of the actual water vapor pressure e to its equilibrium value e^* and is commonly expressed in percent. The tropical mean vertical profile of $\mathcal{R}$ is characterized by large values in the boundary layer, a minimum in the mid troposphere and another maximum in the upper troposphere (Figure 1c). $\mathcal{R}$ also strongly varies regionally within the tropics. Zonal mean $\mathcal{R}$ is largest in the equatorial region, where air rises in the upward branch of the Hadley circulation (Figure 2). The subsiding branches of the circulation in the subtropics, on the other hand, are characterized by minima in zonal mean $\mathcal{R}$ (see also Section 2.3 for an explanation of the dryness of these regions). Locally in these subsidence regions $\mathcal{R}$ is frequently below a few percent (Spencer and Braswell, 1997). Thus, there are large regional variations that are not explained by the Clausius-Clapeyron relation. These are important because, as explained in the following section, the radiative impact of water vapor is controlled by relative rather than by absolute humidity.

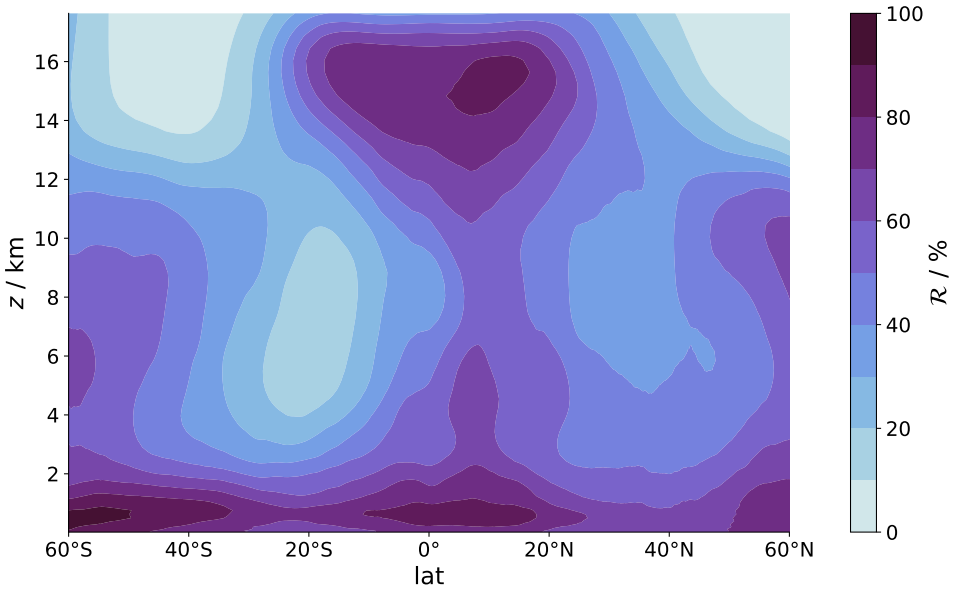


Figure 2: Zonal mean distribution of relative humidity $\mathcal{R}$ as a function of latitude and height in a 45-day (27 June to 11 August) storm-resolving simulation with the ICOsahedral Nonhydrostatic model (ICON).

2.2 IMPORTANCE OF RELATIVE HUMIDITY FOR THE RADIATION BUDGET

2.2.1 *Radiative properties of water vapor*

Water vapor has a strong impact on the Earth's outgoing longwave radiation (OLR). Absorption by water vapor is responsible for about 60% of the total clear-sky greenhouse effect in the thermal infrared (Kiehl and Trenberth, 1997) and reduces the clear-sky OLR (OLR_{cs}) in the tropics by about $100\,W\,m^{-2}$ compared to an atmosphere without water vapor (Pierrehumbert et al., 2007).

The infrared spectrum of water vapor is characterized by distinct absorption bands (e.g. Goody and Robinson, 1951) consisting of thousands of densely packed narrow absorption lines, which correspond to rotational and vibrational modes of molecular excitation. Regions with weak absorption between such bands are called window regions. Within the Earth's emission spectrum there is especially one broad region where water vapor is optically thin, known as the water vapor window, which is located at wavenumbers between about $800\,cm^{-1}$ to $1200\,cm^{-1}$ (Petty, 2006).

Like for most greenhouse gases, the effect of water vapor on OLR_{cs} is approximately logarithmic in specific humidity (e.g. Huang and Shahabadi, 2014; Pierrehumbert, 1999). Due to this logarithmic dependence the radiative impact is determined by fractional rather than absolute changes in water vapor concentration. Small absolute changes in water vapor concentration in the upper troposphere can therefore have a strong radiative impact (e.g. Shine and Sinha,

1991).

2.2.2 *The longwave clear-sky climate feedback*

The humidity of the atmosphere does not only affect the present-day radiation budget, but also the climate feedback. While in my PhD work I focus on the effect of humidity on the present-day radiation budget, the effect on the feedback represents another major motivation to reduce the uncertainty in model-simulated humidity. Therefore, I briefly introduce the concept of the climate feedback, focusing on the longwave clear-sky component, which is directly affected by atmospheric humidity.

In equilibrium the net incoming solar radiation equals the Earth's outgoing longwave radiation such that the net radiation balance at the top of the atmosphere N equals zero. If an external radiative perturbation or forcing is applied to the climate system, such as an increase in greenhouse gas concentrations, N changes by an amount ΔF. In response to this imbalance the surface temperature T_s changes in order to restore equilibrium. Simultaneously, atmospheric and surface processes react in a way that in turn affects N, either by changing the outgoing longwave radiation or the reflected shortwave radiation. Assuming that the effect of all these processes is related to T_s, one can write:

$$\Delta N = \Delta F + \lambda \Delta T_s, \tag{3}$$

where $\lambda = \frac{dN}{dT_s}$ is the climate feedback parameter. The temperature change required to reach a new equilibrium (i.e. $\Delta N = 0$) for a given forcing ΔF is then given by $\Delta T_s^{eq} = -\Delta F / \lambda$.

The total feedback parameter λ can be divided into several components. The most fundamental feedback is the Planck feedback. It describes the change in outgoing longwave radiation that results from a vertically uniform temperature change only: in response to a radiation imbalance the temperature of the Earth's surface and the overlying atmosphere increase until a new radiative equilibrium is reached. This represents a strong negative (stabilizing) feedback. In the real atmosphere, however, deviations from this basic Planck response cause additional feedbacks. The two most important feedbacks acting in the longwave spectral region in the absence of clouds are the water-vapor and lapse-rate feedbacks. Together with the Planck feedback they constitute the longwave clear-sky feedback λ_{cs}, which is defined as:

$$\lambda_{cs} = \frac{d\,OLR_{cs}}{d\,T_s}. \tag{4}$$

The water-vapor feedback results from increasing water vapor concentrations with warming, which leads to an upward shift of the emission height and thereby diminishes longwave emission compared to the Planck response. In consequence, a stronger surface warming is required to restore equilibrium,

making the water-vapor feedback a positive feedback. The lapse-rate feedback results from a vertically non-uniform warming of the atmosphere. In the tropics and mid-latitudes the upper troposphere warms faster than the surface, enabling the Earth to radiate to space more efficiently. The lapse-rate feedback is therefore negative and partly offsets the positive water vapor feedback. Due to their strong anti-correlation water vapor and lapse-rate feedbacks are typically considered jointly as the combined water-vapor–lapse-rate feedback (e.g. Bony et al., 2006; Soden and Held, 2006).

Apart from the longwave clear-sky feedbacks described here, other important climate feedbacks are caused by changes in clouds and in surface albedo. I do not further elaborate on these here because they are not directly affected by atmospheric humidity.

2.2.3 *The partly-Simpsonian nature of the atmosphere and the importance of relative humidity*

Much of the response of OLR_{cs} to changes in surface temperature, atmospheric temperature structure and water vapor concentration can be understood by dividing the spectrum of OLR_{cs} into two regions: the optically thin water vapor window and the optically thick water vapor bands. Within the window, the radiation emitted to space originates from the Earth's surface and therefore depends only on surface temperature according to Planck's law. In the water vapor bands, the emitted radiation originates from within the atmosphere, roughly from the altitude region at which the optical thickness τ reaches a value of 1 at the respective wavenumber (e.g. Petty, 2006). The emission from the water vapor bands is therefore independent of surface temperature. To first order it only depends on $\mathcal{R}$ (e.g. Nakajima et al., 1992). As shown by Jeevanjee et al., 2021, if temperature and water vapor concentration jointly change in such a way that $\mathcal{R}$ remains unchanged, the effects induced by changes in thermal emission and by the shift in emission height cancel almost completely. This behaviour was first articulated by Simpson, 1928 and has therefore been referred to as Simpsonian (e.g. Ingram, 2010; Seeley and Jeevanjee, 2021).

The partly-Simpsonian behavior of the atmosphere has important consequences for OLR_{cs} and in turn for the feedback λ_{cs}. One important consequence is the positive combined water-vapor–lapse-rate feedback. If the atmosphere warms at roughly constant $\mathcal{R}$, as predicted by climate models and theory (e.g. Held and Soden, 2000; Romps, 2014), the emission to space changes little within the water vapor bands. The emission can only increase in the window regions of the spectrum (e.g. Koll and Cronin, 2018), which strongly reduces the efficiency at which the Earth emits energy to space compared to the Planck response and therefore represents a positive feedback.

The partly-Simpsonian model also explains why small deviations from a constant $\mathcal{R}$ under warming strongly impact λ_{cs}. The emission from the water vapor bands primarily depends on $\mathcal{R}$. Thus, a change of $\mathcal{R}$ under warming allows OLR_{cs} to change in the water vapor bands and thereby strongly amplify or weaken the feedback compared to a warming at constant $\mathcal{R}$. Since the relation between OLR_{cs} and $\mathcal{R}$ is non-linear, OLR_{cs} is particularly sensitive to changes in $\mathcal{R}$ in the dry tropical subsidence regions (Spencer and Braswell, 1997). To date it is not known how exactly tropical $\mathcal{R}$ will change under warming. GCMs on average predict a slight reduction of $\mathcal{R}$ in the subsidence regions (Sherwood et al., 2010), which reduces the global water vapor feedback by $\sim$5 % compared to the feedback resulting from the assumption of constant $\mathcal{R}$ (Soden and Held, 2006). However, there is a considerable spread in the $\mathcal{R}$ response across models, which drives the current model spread in λ_{cs} (Vial et al., 2013).

Another mechanism by which $\mathcal{R}$ can affect OLR_{cs} and hence λ_{cs} is by closing the water vapor window, i.e. by increasing the portion of the emission spectrum that behaves in a Simpsonian manner. The optical thickness in the window region depends on the water vapor path of the atmosphere. At high surface temperatures of about 300 K and correspondingly high water vapor concentrations, the optical thickness sufficiently increases to a level such that a further addition of water vapor rapidly closes the window (Koll and Cronin, 2018). For a given tropical temperature profile, variations in $\mathcal{R}$ therefore significantly affect the spectral width of the window (McKim et al., 2021) and hence OLR_{cs}. It has been shown that due to this mechanism the tropical λ_{cs} does not only depend on the $\mathcal{R}$ response to warming, but also on present-day $\mathcal{R}$ (Bourdin et al., 2021; Kluft et al., 2019; McKim et al., 2021).

The $\mathcal{R}$-dependence underlines the importance of the dry tropical subsidence regions. Not only are they particularly sensitive to changes in $\mathcal{R}$ under warming, but even if $\mathcal{R}$ is assumed to stay unchanged, λ_{cs} is most negative in these regions (McKim et al., 2021). Due to their importance in stabilizing Earth's climate they have been referred to as the "radiator fins" of the tropics (Pierrehumbert, 1995).

Due to the central role of $\mathcal{R}$ in determining OLR_{cs} and λ_{cs}, I focus on inter-model differences in tropial $\mathcal{R}$ rather than specific humidity in my PhD work. From the above, it is clear that the radiative effect of these differences can only be determined based on the entire distribution of $\mathcal{R}$ rather than the tropical mean. The following section introduces a simple model of how this distribution is established.

2.3 A SIMPLE MODEL FOR THE DISTRIBUTION OF RELATIVE HUMIDITY – THE LAST-SATURATION PARADIGM

The processes controlling the humidity distribution in the tropical free troposphere are complex, as they include advective processes on all scales, ranging from synoptic-scale over convective to turbulent mixing, as well as advection and

evaporation of cloud particles and precipitation (Emanuel and Pierrehumbert, 1996). In the last two decades, great progress in understanding the distribution of free-tropospheric $\mathcal{R}$ has been made by implementing the last-saturation or advection-condensation paradigm (Pierrehumbert et al., 2007). This framework only considers advection and saturation and thus represents the simplest possible perspective on how the $\mathcal{R}$ distribution is established. In the second study presented in this thesis, I use the last-saturation paradigm as a null-hypothesis to investigate the physical mechanisms behind differences in relative humidity between simulations.

The last-saturation paradigm is based on the assumption that the specific humidity q of an air parcel behaves as a conservative tracer, for which condensation is the only sink. Under this assumption, the specific humidity of the air parcel is conserved after the point at which it was last saturated (i.e. after leaving a cloud), so that the specific humidity at any given target point $\mathcal{R}_t$ is equal to the saturation specific humidity at last saturation q^*_{ls}:

$$\mathcal{R}_t = \frac{q_t}{q^*_t} = \frac{q^*_{ls}}{q^*_t} \approx \frac{e^*(T_{ls})}{e^*(T_t)} \frac{p_t}{p_{ls}}, \tag{5}$$

where the indices t and ls denote the target point and the point of last-saturation, respectively. Here, $\mathcal{R}$ is defined as the ratio of q to its saturation value q^*. While this definition is less common than that based on water vapor pressure (see Section 2.1), it is more convenient within the last-saturation framework because q is a materially conserved quantity. Numerical values yielded by the two definitions differ by less than 1 %. To obtain the last expression in Equation 5, q is approximated according to Equation 2. According to Equation 5 $\mathcal{R}_t$ is simply the fraction of saturation water vapor pressures e^* at two temperatures, multiplied by a pressure ratio. In practice, changes in last-saturation pressure have a minor impact, which is why for a given target pressure $\mathcal{R}_t$ primarily depends on last-saturation and target temperature. Due to their important role in setting the humidity distribution within the last-saturation framework, the atmospheric regions where last-saturation typically occurs are also referred to as the "source regions" or "origins" of the air arriving in a given target region. The last-saturation paradigm neglects any processes that could affect an air parcel's q after last-saturation. These processes include evaporation or sublimation of hydrometeors advected with or sedimenting through the air parcel, as well as mixing with air parcels of different humidity.

The evolution of $\mathcal{R}$ during the typical life cycle of a tropical air parcel is schematically illustrated in Figure 3. In the boundary layer the parcel is supplied with moisture by evaporation from the surface and therefore has a high $\mathcal{R}$. As it rises in a convective updraft it cools, which causes $\mathcal{R}$ to increase. As soon as saturation is reached, the q of the parcel decreases due to condensation. At some altitude level the parcel stops its ascent, detrains from the cloud and moves through the atmosphere horizontally. During this horizontal advection the air parcel looses energy by radiative cooling, causing it to gradually subside. While it subsides the air parcel is compressed and its temperature increases, while it

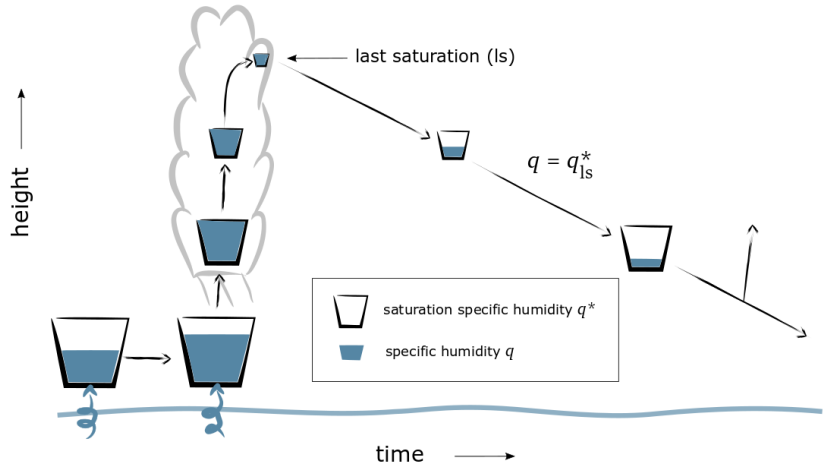


Figure 3: Illustration of the typical life of a tropical air parcel and the evolution of its relative humidity according to the last-saturation paradigm. The air parcel is depicted as a bucket with a size representing saturation specific humidity q^*, and a content representing specific humidity q. The fill level of the bucket hence represents its relative humidity. The last-saturation paradigm assumes that the specific humidity of the air parcel is conserved after the point of last saturation.

maintains the q it had when it was last saturated. Hence, its $\mathcal{R}$ decreases. The quasi-horizontal transport continues until the parcel is either entrained into another convective system or it reaches the boundary layer again. This simple picture already explains how an air parcel can end up with very low $\mathcal{R}$ in the free troposphere.

If the atmospheric circulation and temperature distribution are known, regional $\mathcal{R}$ variations within the tropics can be qualitatively understood from the last-saturation paradigm. Figure 4 illustrates this based on the mid-tropospheric $\mathcal{R}$ (4 km to 8 km averaged) field from a 25-day global storm-resolving simulation and corresponding last-saturation statistics obtained from back-trajectories (both the simulation and the trajectory calculations are described in more detail in Section 5). To demonstrate how moist and dry tropical regions differ in their last-saturation statistics, two exemplary target regions are selected: a region of very high $\mathcal{R}$ in the Eastern Equatorial Pacific and a region of very low $\mathcal{R}$ in the Eastern South Atlantic. The last-saturation events associated with the moist region are located very close to the target region both in the horizontal (Figure 4a) and in the vertical (Figure 4b). The corresponding parcels have therefore experienced only a weak subsidence drying since last-saturation or are still saturated. In other words, regions of high $\mathcal{R}$ are located in or close to deep convective or mid-level clouds, which are concentrated in the Intertropical Convergence Zone (ITCZ). In contrast, the air parcels associated with the dry target region have been advected over larger horizontal and vertical distances since last-saturation. During their

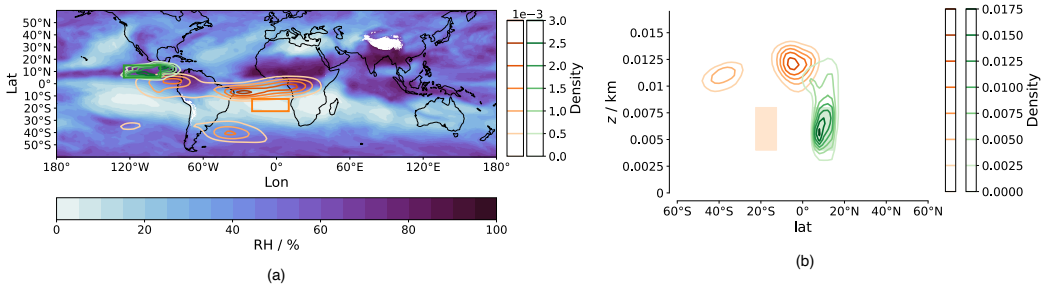


Figure 4: Origins of air parcels arriving in two exemplary regions of high and low relative humidity. (a) Mid-tropospheric (4 km to 8 km vertically averaged) relative humidity $\mathcal{R}$ averaged over a 25-day (17 July to 11 August) period simulated by the ICON storm-resolving model (shading), as well as number densities of last-saturation points (contours) associated with two exemplary target regions (rectangles) of high $\mathcal{R}$ (green) and low $\mathcal{R}$ (orange). (b) Number densities of last-saturation points as a function of latitude and height (contours). Rectangles shaded in green and orange indicate the moist and dry target regions, respectively.

descent from the upper troposphere the parcels warm substantially, resulting in a low $\mathcal{R}$ (Equation 5).

For the dry zone selected in Figure 4, last-saturation events are located both in the equatorial region and in the extra-tropics. This is not an exception, as a significant part of the dry subtropical air originates from outside the tropics (e.g. Roca et al., 2012). Air parcels originating from the inner tropics typically experience last-saturation when detraining from deep convective clouds in the upper troposphere (∼12 km) and subsequently gradually descend across isentropic surfaces due to radiative cooling (as described above for the tropical air parcel illustrated in Figure 3). In the extra-tropics, the primary dehydration process is isentropic mixing: air that is transported from the tropics into mid-latitudes rises along upward sloping isentropes, saturates there, and then descents back into the tropics following isentropes (e.g. Galewsky et al., 2005). In both cases, air subsides undisturbed over a large vertical distance and thereby experiences drying.

To test whether the last-saturation framework quantitatively explains tropical $\mathcal{R}$ variations, a number of studies have implemented it numerically, using large-scale wind and temperature fields from meteorological analyses to calculate Lagrangian back-trajectories (e.g. Dessler and Sherwood, 2000; Pierrehumbert and Roca, 1998; Sherwood, 1996). None of these implementations explicitly accounted for evaporation of hydrometeors or any mixing by motions on scales smaller than those incorporated in the trajectories. Despite that, they successfully reproduced the observed large-scale tropical humidity field, typically within an accuracy of about 10%. This has lead to the conclusion that, if the large-scale atmospheric circulation and temperature structure are known, any moistening or drying from microphysical processes or unresolved mixing play a secondary role in establishing the tropical humidity distribution. This is not to say that these

processes are unimportant, rather to say that to the extent they are important, it is through their indirect influence on the atmospheric circulation and the temperature structure, which ultimately determine the location of last-saturation events. Moisture sources or sinks caused by these processes after last-saturation play a secondary role in explaining the large-scale $\mathcal{R}$ distribution. It is, however, less clear whether these sources and sinks might be important when it comes to explaining more subtle humidity differences between models, in particular when different parameterizations are used to represent microphycis and turbulence. This is what I test in my second study.

3

OBSERVATIONS OF HUMIDITY AND CLOUDS IN THE TROPICS (ACCOMPANYING PROJECTS)

In this chapter I briefly summarize two accompanying projects I pursued during my PhD term and that resulted in publications or co-authorships, respectively.

3.1 AN OBSERVATIONAL CLIMATE DATA RECORD OF FREE-TROPOSPHERIC HUMIDITY

As described in Section 2.2, the emission to space within the spectral regions dominated by water vapor absorption strongly depends on atmospheric relative humidity. Inversely, this can be used to retrieve relative humidity from measurements by passive infrared or microwave sensors onboard satellites (e.g. Buehler, 2005; Soden and Bretherton, 1993). These so-called humidity sounders measure the average relative humidity over broad layers in the free troposphere. Operational weather satellites have carried such instruments over several decades. This offers the possibility to create long-term data records, which can be used to track the evolution of free-tropospheric humidity and validate climate models (e.g. Buehler et al., 2008; Shi and Bates, 2011). In my master thesis project I created a new climate data record (CDR) of upper-tropospheric humidity (UTH) based on observations from three different microwave humidity sounders on eleven satellites. During my PhD period, I performed a comparison of the new CDR to an existing data record and published a description of the CDR including the comparison:

Lang, T., Buehler, S. A., Burgdorf, M., Hans, I. and John, V. O. (2020). A new climate data record of upper-tropospheric humidity from microwave observations. *Scientific Data*, 7, 218. https://doi.org/10.1038/s41597-020-0560-1.

A major novelty of this CDR is that a different definition of UTH is used in the retrieval from the brightness temperatures measured by the microwave sensor. The same definition can be used to derive UTH from infrared observations, allowing a more synergistic use of infrared- and microwave-derived UTH in the future. Furthermore, the CDR is based on a lower-level data record that uses a new approach for re-calibration of brightness temperatures. My comparison shows that this results in a reduction of inter-satellite biases compared to the existing data record.

Recently, the UTH CDR was part of a larger intercomparison with three other microwave and infrared UTH data records, for which I provided combined time series of UTH with measurements from overlapping satellite missions. A manuscript on this intercomparison, which I co-authored, has recently been

accepted for publication:

Shi, L., Schreck III, C. J., John, V. O., Chung, E.-S., Lang, T., Buehler, S. A., and Soden, B. J. Assessing the consistency of satellite derived upper tropospheric humidity measurements, *Atmospheric Measurement Techniques Discussion* [preprint]. https://doi.org/10.5194/amt-2022-204

The main outcome of the intercomparison is that all four data sets show a similar seasonal variability of UTH, but there are significant disagreements on the decadal trends of UTH across the data sets.

3.2 OBSERVATIONS OF SHALLOW CLOUDS DURING THE EUREC^{4}A FIELD CAMPAIGN

Due to their dryness, the tropical subsidence regions have a profound impact on the longwave clear-sky feedback (see Section 2.2). At the same time, the subsidence regions also play a key role in determining the shortwave component of the climate feedback: The descending air creates a temperature inversion in the lower troposphere, which prevents clouds from growing deep. The subsidence regions are therefore populated by shallow clouds, of which trade-wind cumuli are the most frequently occurring type. These clouds only weakly affect OLR, because their temperature is not significantly different from the surface. However, they efficiently reflect shortwave solar radiation and thereby influence the Earth's radiation budget. How these clouds respond to warming is therefore critical for the cloud feedback, and model differences in this response explain the largest part of the current spread of climate sensitivity across GCMs (e.g. Bony, 2005; Medeiros et al., 2014; Vial et al., 2013). Differences in the representation of lower-tropospheric vertical mixing associated with convection and large-scale circulations are thought to be the cause of this inter-model spread (e.g. Sherwood et al., 2014; Vial et al., 2016).

To gain a better understanding of the interplay between clouds and circulation, the field campaign *Elucidating the role of clouds-circulation coupling in climate* (EUREC^{4}A) was launched in early 2020. It took place in the Atlantic trade-wind regime, over the ocean east of Barbados, and included aircraft measurements as well as surface and ship-based observations. As part of my PhD project I actively participated in the EUREC^{4}A campaign. During the two weeks I spent on site on Barbados, one of my main tasks was to launch radiosondes from the Barbados Cloud Observatory (BCO). Besides that, I had the chance to accompany one of the research flights with the High Altitude and Long Range Research Aircraft (HALO; Krautstrunk and Giez, 2012) and launch dropsondes from the aircraft. I also assisted the ground control for the HALO aircraft. Furthermore, our group was responsible for compiling a daily weather report, for which I analyzed radiosonde data. Due to these contributions I am a co-author of the campaign overview paper:

Stevens, B., Bony, S., Farrell, D., and 290 co-authors (2021). EUREC4A, *Earth System Science Data*, 13, 4067–4119. https://doi.org/10.5194/essd-13-4067-2021

After the campaign, I joined the HALO flight phase segmentation initiative, which aimed at promoting a consistent analysis of the flight data by providing metadata on different modes of operation of the aircraft during a flight. To this end, we divided each of the 15 research flights into several standard elements or "flight segments", each of which is defined by a start and end time and belongs to a pre-defined "kind" that indicates the operation mode of the aircraft. This allows to distinguish between the circular tracks HALO followed most of the time, straight excursions from the circle and different instrument calibration maneuvers. Moreover, with each segment the IDs of dropsondes that were launched within the respective time period are provided. This enables a convenient selection of sondes from the EUREC^{4}A dropsonde data set (JOANNE; George et al., 2021) based on flight segments, which has been used in the generation of circle-averaged data products (e.g. Albright et al., 2022; George et al., 2021). The flight segmentation data are provided as YAML (YAML Ain't Markup Language) files and are openly accessible online (https://doi.org/10.5281/zenodo.4900003). My contribution lead to a co-authorship of the paper describing the EUREC^{4}A HALO data:

Konow, H., Ewald, F., George, G., and 35 co-authors (2021). EUREC4A's HALO, *Earth System Science Data*, 13, 5545–5563. https://doi.org/10.5194/essd-13-5545-2021

4

QUANTIFYING THE HUMIDITY SPREAD AND ITS IMPACT ON THE RADIATION BUDGET ACROSS GLOBAL STORM-RESOLVING MODELS (PAPER 1)

Reducing the model spread in tropical free-tropospheric relative humidity ($\mathcal{R}$) and its response to warming is a crucial step toward reducing the uncertainty in clear-sky climate sensitivity. This is hoped to be achieved with recently developed global storm-resolving models (GSRMs). The study included in Appendix A is the first to quantify the tropical $\mathcal{R}$ spread in this new generation of models, making use of the first multi-model ensemble of GSRMs. It addresses the following research questions:

1. **How large is the spread in tropical relative humidity across global storm-resolving models?**
2. **How relevant is the spread in terms of its impact on the clear-sky radiation budget?**
3. **In which regions of the tropical atmosphere is the spread most relevant?**

In the following, I summarize the main methods I applied to answer these questions and the key results I obtained.

4.1 THE HUMIDITY SPREAD IN THE DYAMOND ENSEMBLE

The research questions are addressed by comparing nine different models that participated in the Atmospheric general circulation Modeled On Nonhydrostatic Domains (DYAMOND) project (Stevens et al., 2019). As an observationally constrained reference data set I include the ERA5 reanalysis (Hersbach et al., 2020) in the comparison. The DYAMOND models differ in many aspects, ranging from the type of numerical grid they use to solve their governing equations to their parameterizations of subgrid processes. This diversity makes the DYAMOND ensemble well-suited to assess the dependence of $\mathcal{R}$ on model formulation. The simulations cover a 40-day period in northern-hemisphere summer. After a common initialisation the atmosphere was allowed to evolve freely without further forcing. Daily observed sea-surface temperature and sea-ice concentrations were used as boundary conditions. For the analysis I exclude the first ten days of the simulations to minimize constraints from the common initialization as well as biases due to differences in model spin-up. The size of the model output ($\sim$60 TB) represents a challenge for the analysis. By performing a random subsampling of grid points I reduce the amount of data by about a factor of 100, while ensuring that the resulting sampling error is negligible compared to inter-model differences.

I find that the DYAMOND models all capture the typical C-shape of the tropical mean vertical profile of $\mathcal{R}$ (Figure 5a) and the multi-model mean agrees remarkably well (within 2% of $\mathcal{R}$) with the ERA5 profile. This suggests that there is at least no systematic moist or dry bias of $\mathcal{R}$ with respect to observations that is shared by all GSRMs, as it is the case for GCMs (e.g. John and Soden, 2007; Pierce et al., 2006). This should be verified with other observational data sets, since ERA5 itself is known to exhibit a moist bias with respect to satellite observations, particularly in dry regions (Xue et al., 2020).

$\mathcal{R}$ differences among the models, however, are considerable. The inter-model standard deviation of $\mathcal{R}$ ($\sigma(\mathcal{R})$) exceeds 5% in the upper troposphere as well as in the lower troposphere around the top of the boundary layer (Figure 5c). The latter is related to model differences in boundary layer depth. As the transition from the boundary layer to the free troposphere is marked by a steep gradient in $\mathcal{R}$, differences in boundary layer depth cause a large inter-model spread in $\mathcal{R}$. In the mid troposphere the $\mathcal{R}$ spread is smallest ($\sigma(\mathcal{R}) \approx 2\,\%$).

To examine how the $\mathcal{R}$ spread in the DYAMOND GSRMs compares to that in conventional atmospheric GCMs, I 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). This comparison shows that in DYAMOND, the spread in tropical mean $\mathcal{R}$ is reduced throughout most of the free troposphere, except from the tropopause region and the transition to the boundary layer (Figure 5c). The relative reduction is largest in the upper troposphere ($\sim$50 % to 70 %) and smaller in the mid troposphere ($\sim$25 % to 50 %). The reduction indicates that $\mathcal{R}$ is indeed better constrained in GSRMs than in GCMs. The reduced spread is a promising result, especially considering that the models were not specifically tuned for the DYAMOND simulations and some of them even ran for the first time in this particular configuration and at storm-resolving resolutions.

A legitimate question that arises is how much one can learn about climatological $\mathcal{R}$ biases from an intercomparison as short as 30 days. One important concern is that the identified $\mathcal{R}$ differences might reflect a poor sampling of internal variability. However, the DYAMOND inter-model spread is significantly larger than what would be expected from internal variability, which is estimated from five years of ERA5 reanalysis data (Figure 5c). This indicates that the inter-model $\mathcal{R}$ differences mostly represent systematic differences due to differences in model formulation. While this is true for the tropical mean, I show in my second study that $\mathcal{R}$ differences in the driest tropical regions are indeed strongly coloured by internal variability (see Section 5.2).

Besides the comparison of tropical means, I also compare the $\mathcal{R}$ distribution in moisture space (e.g. Bretherton et al., 2005; Schulz and Stevens, 2018). Moisture space is spanned by sorting the atmospheric state from dry to moist based on the vertically integrated amount of water vapor (IWV). The IWV-ranked profiles are

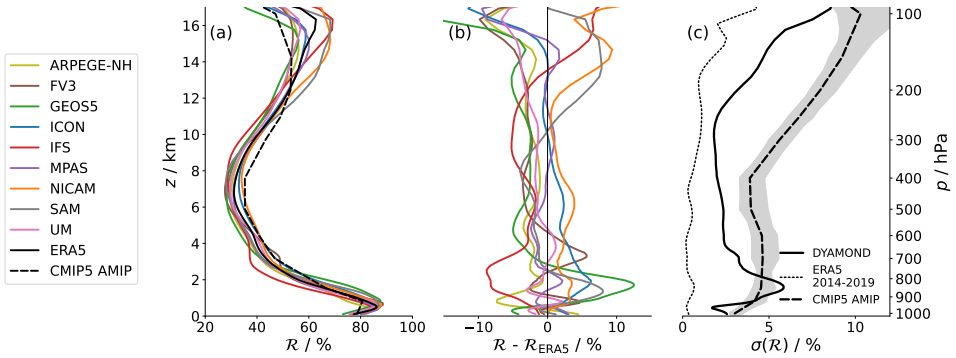


Figure 5: Tropical mean profiles of relative humidity $\mathcal{R}$ and inter-model spread in the DYAMOND ensemble. (a) Tropical mean vertical profiles of $\mathcal{R}$ over ocean regions from all nine DYAMOND models (colours), the ERA5 reanalysis (black solid) and the CMIP5 AMIP 30-year multi-model mean (black dashed). (b) Vertical $\mathcal{R}$ profiles for the DYAMOND models shown as deviation from the ERA5 profile. (c) Inter-model standard deviation of tropical mean $\mathcal{R}$ in DYAMOND (solid line). For comparison, the inter-annual $\mathcal{R}$ spread in five years of ERA5 (2014-2019; dotted line) as well as the inter-model spread of the 30-year mean $\mathcal{R}$ 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.

aggregated within 50 equal-sized blocks and averaged. This allows to distinguish between different dynamic regimes of the tropics: deep convective regimes are associated with high IWV percentiles, subsidence regimes with low IWV percentiles. The representation in moisture space reveals that the large inter-model spread in the upper troposphere and around the top of the boundary layer ($\sigma(\mathcal{R}) > 5\,\%$) prevails throughout the tropics (Figure 6a). In the mid troposphere, however, the spread increases from dry to moist regimes and maximizes between the 80th to 90th percentile of IWV ($\sigma(\mathcal{R}) > 5\,\%$). This region of maximum spread roughly marks the transition from deep convective to subsidence regimes in moisture space and is associated with upper-level anvil clouds, which could point to differences in the representation of convection, small-scale mixing and/or microphysical processes as possible sources.

4.2 THE IMPACT OF HUMIDITY DIFFERENCES ON OUTGOING LONGWAVE RADIATION

To assess the relevance of the $\mathcal{R}$ differences between the DYAMOND models, I translate them into differences in clear-sky outgoing longwave radiation (OLR_{cs}) using a fast radiative transfer model. The calculations are performed based on the block-averaged atmospheric state in moisture space rather than based on individual instantaneous profiles. This strongly reduces the computing effort, while still capturing most of the non-linearity of OLR_{cs} in $\mathcal{R}$ (see Section 2.2). Using the same radiation scheme for each model and fixing the concentrations of other greenhouse gases than water vapor ensures that the resulting differences in OLR_{cs} solely reflect differences in atmospheric temperature and water vapor

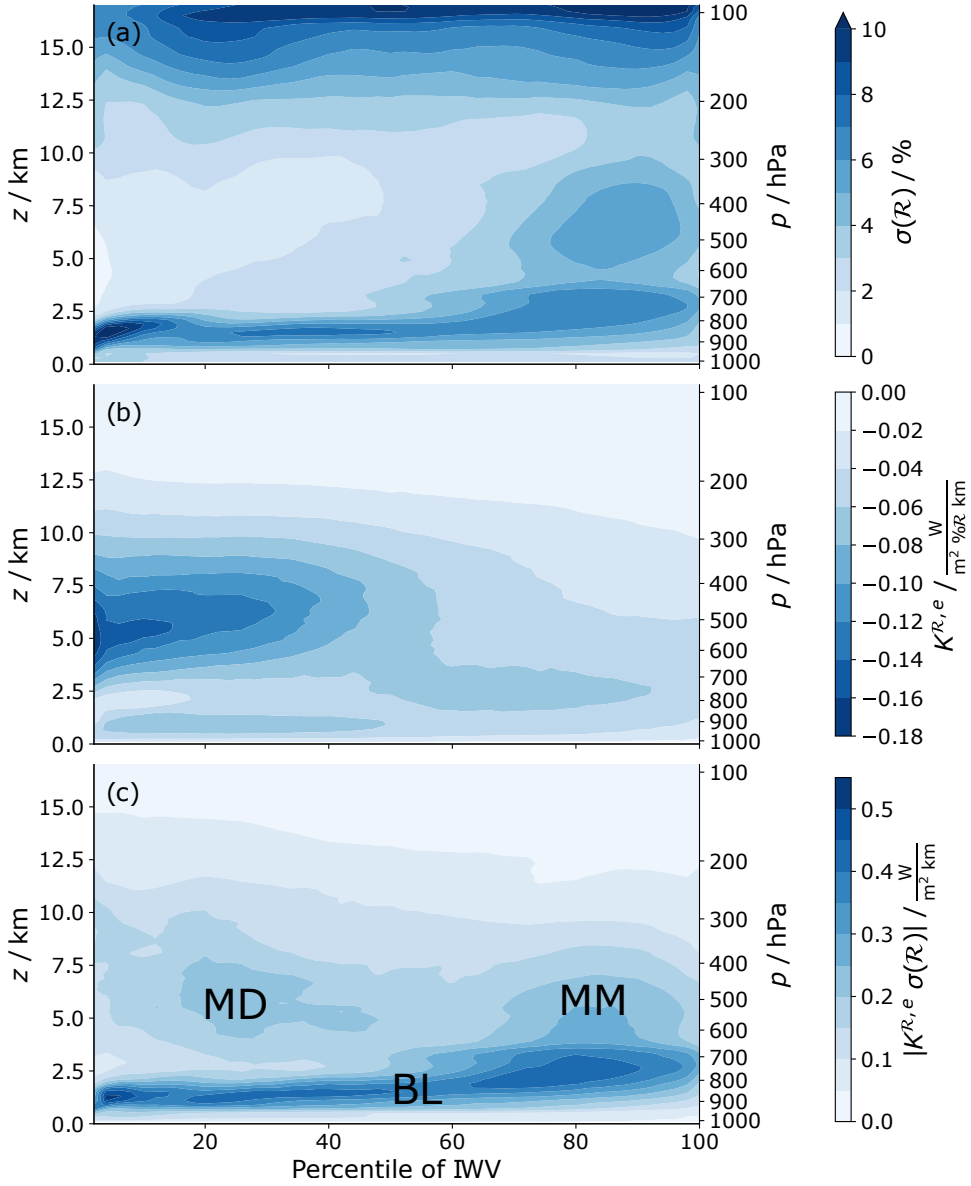


Figure 6: Inter-model spread in relative humidity $\mathcal{R}$ across the DYAMOND models and its effect on clear-sky outgoing longwave radiation (OLR_{cs}) in moisture space. (a) Inter-model standard deviation of $\mathcal{R}$ ($\sigma(\mathcal{R})$) as a function of height and percentile of vertically integrated water vapor (IWV). (b) Relative humidity response kernel ($K^{\mathcal{R},e}$) showing the sensitivity of OLR_{cs} to a 1 % (additive) change in $\mathcal{R}$ in a 1 km layer under constant temperature. (c) Inter-model standard deviation $\sigma(\mathcal{R})$ weighted with $K^{\mathcal{R},e}$. Acronyms denote three regions of particular relevance: The transition between the boundary layer and the free troposphere (BL), the mid troposphere of moist regimes (MM) and the mid troposphere of dry regimes (MD).

concentrations. The standard deviation in tropical mean OLR_{cs} is $1.2\,W\,m^{-2}$. The most extreme models differ by about $4\,W\,m^{-2}$. These differences are of similar magnitude as the radiative forcing from a doubling of CO_2 ($3.7\,W\,m^{-2}$) and hence relevant.

To identify the regions in which a reduction in $\mathcal{R}$ spread would be most beneficial, I use a radiative kernel technique to determine the impact of $\mathcal{R}$ differences on OLR_{cs} in different altitude layers and different regimes in moisture space. The relative humidity kernel describes the sensitivity of OLR_{cs} to perturbations in $\mathcal{R}$ and is calculated by successively applying small perturbations in water vapor pressure e in different altitude regions, using ERA5 as the base state. (Since $\mathcal{R}$ is a function of vapor pressure and temperature, another possible way to determine the $\mathcal{R}$ kernel is be to perturb the temperature instead of vapor pressure. I elaborate more on the differences between these two approaches in Sections A.4.3 and A.6.2.) The $\mathcal{R}$ sensitivity kernel is characterised by a distinct maximum in the mid troposphere of the driest tropical regimes (Figure 6b), which is in qualitative agreement with earlier studies (e.g. Spencer and Braswell, 1997).

Weighting the $\mathcal{R}$ kernel with the inter-model spread in $\mathcal{R}$ reveals how much model differences contribute to differences in OLR_{cs} within a certain region (Figure 6c). I identify three regions of particular relevance (marked by acronyms in Figure 6c). One is the transition between the boundary layer and the free troposphere (BL). In this region model differences in $\mathcal{R}$ are large and therefore cause a significant part (about one third) of the spread in OLR_{cs}, although the sensitivity of OLR_{cs} to a given $\mathcal{R}$ perturbation is low in the lower troposphere. The remaining part (about two thirds) of the OLR_{cs} spread is caused by $\mathcal{R}$ differences in the mid troposphere. The combination of increasing $\mathcal{R}$ spread and decreasing sensitivity of OLR_{cs} from dry to moist regions results in a relatively uniform importance of $\mathcal{R}$ differences across the tropics. Local maxima occur in the mid troposphere of dry regions (MD) near the 30th percentile of IWV and in the mid troposphere of moist regions (MM) near the 80th IWV percentile. The impact of the comparably large $\mathcal{R}$ differences in the upper troposphere is negligible, because there OLR_{cs} is virtually insensitive to changes in $\mathcal{R}$.

These results are limited to the clear-sky case. High clouds, which are most abundant in the moist regimes, would mask some of the clear-sky effect and hence dampen the impact of the large $\mathcal{R}$ differences in the high IWV percentiles. This emphasizes even more the importance of the dry regimes, where high clouds are rare.

In summary, the first study yields the following key results:

1. The spread in tropical free-tropospheric relative humidity across global atmospheric storm-resolving models is approximately halved compared to that across general circulation models.

2. The remaining humidity differences still cause a relevant ($\sim 1.2\,W\,m^{-2}$) spread in tropical mean clear-sky OLR.

3. A further reduction of humidity spread in the lower and mid free troposphere, both in dry subsidence regimes and in moist regimes at the transition to deep convective regimes, would be most beneficial to reduce the spread in clear-sky OLR.

5

UNDERSTANDING THE SOURCES OF HUMIDITY SPREAD ACROSS GLOBAL STORM-RESOLVING MODELS (PAPER 2)

The analysis summarized in the previous chapter showed that humidity differences across GSRMs are still relevant in terms of their radiative impact. This motivated the study included in Appendix B, in which I aim to better understand the sources of uncertainties in modelling processes that drive the distribution of humidity. I approach this by asking two research questions:

1. **How much do different model uncertainties contribute to the spread in relative humidity across global storm-resolving models?**
2. **Through which physical mechanisms do these model uncertainties affect relative humidity?**

5.1 SENSITIVITY OF RELATIVE HUMIDITY TO DIFFERENT MODEL UNCERTAINTIES

To answer the first question, I run a series of experiments with a GSRM. In these I examine how much relative humidity $\mathcal{R}$ in the tropics changes in response to different modifications in model formulation. These modifications are inspired by the differences across the DYAMOND models (see Section 4) and thus represent differences in current modeling approaches across modeling groups. I investigate uncertainties in both the dynamical core and model physics by applying changes to model resolution and parameterizations.

The experiments are performed with the ICOsahedral Nonhydrostatic model (ICON; Zängl et al., 2015) in its storm-resolving setup (Hohenegger et al., 2022). The control experiment is run with a horizontal grid spacing of $\sim$5 km and 110 terrain-following vertical levels, which corresponds to a vertical grid spacing of 400 m in the mid to upper troposphere over an ocean surface. In the control experiment turbulent mixing is represented by a 3D Smagorinsky-Lilly scheme, microphysical processes are represented by a one-moment bulk scheme. The simulated time period spans 45 days in June to August 2021 and the experimental protocol closely follows that specified by the DYAMOND intercomparison (see Section B.3). To test to what extent the humidity differences in the 45-day experiments might reflect sampling error, I perform an additional control experiment (Control 2) with perturbed initial conditions.

Three experiments test the sensitivity of $\mathcal{R}$ to changes in model resolution: In the "$\Delta x/2$" experiment the horizontal grid spacing is halved relative to the control experiment to 2.5 km. In the "$2\Delta z$" and "$\Delta z/2$" experiments the number of vertical levels is changed such that the vertical grid spacing in the free troposphere is

increased to 800 m and decreased to 200 m, respectively. Three further experiments examine how $\mathcal{R}$ is affected by changes in model parameterizations. Here, I focus on contrasting schemes, which I consider a stronger perturbation than perturbing single parameters within a given scheme. For the "TTE" experiment the 3D Lilly-Smagorinsky turbulence scheme is exchanged by a 1D total turbulent energy (TTE) scheme and in the "2-mom" experiment the one-moment microphysics scheme is exchanged by a two-moment scheme. Additionally, I also explore a common parameter sensitivity by increasing (approximately doubling) the terminal fall speed of ice particles in the $2v_{\text{ice}}$ experiment. For more background on these choices the reader is referred to Section B.2.

The sensitivity experiments produce a spread in tropical mean $\mathcal{R}$ that is similar to the inter-model spread in the DYAMOND ensemble found in the first study (Figure 7c). There, I showed that this spread is smaller than that across classical GCMs, which indicates that resolving deep convection reduces the model spread in tropical $\mathcal{R}$. The experiments performed in this second study support this by showing that even strong perturbations in a GSRM cannot reproduce the spread in $\mathcal{R}$ seen across models with convective parameterizations. Perturbing the convective parameterization in a GCM, however, does reproduce a major part of the inter-model spread of tropical clear-sky OLR (which largely represents differences in $\mathcal{R}$) across GCMs (Tsushima et al., 2020).

With few exceptions the changes in tropical mean $\mathcal{R}$ that are produced in the sensitivity experiments are larger than the difference between the two control experiments (Figure 7a, b), which indicates that they represent systematic changes rather than internal variability. Halving the horizontal grid spacing ($\Delta x/2$) and halving the vertical grid spacing ($\Delta z/2$) lead to modest ($\leq$3 %) changes in tropical mean $\mathcal{R}$, whereas doubling the vertical grid spacing ($2\Delta z$) results in a strong (>10 %) $\mathcal{R}$ increase in the upper troposphere. The latter experiment, however, represents a rather extreme case, in the sense that GSRMs are not commonly run at such coarse vertical grid spacings. Therefore, these results suggest that $\mathcal{R}$ is rather robust to changes in model resolution within the general scale of GSRM resolution. This does not exclude the possibility that reducing the horizontal grid spacing to much finer scales (on the order of 200 m) could make a difference, which needs to be tested in future experiments.

Larger $\mathcal{R}$ changes occur in the experiments with exchanged microphysics and turbulence parameterizations. They maximize in the lower and mid troposphere, where they have the largest impact on clear-sky OLR, as I showed in the first study. Changing the turbulence scheme results in a $\mathcal{R}$ increase of about 8 % in a broad mid-tropospheric layer. Employing the 2-mom microphysics scheme reduces $\mathcal{R}$ by about 10 % in a rather thin lower-tropospheric layer, reflecting a shallowing of the boundary layer. This could be related to an earlier onset of precipitation in the 2-mom experiment, which results in clouds growing less deep (Stevens and Seifert, 2008). In summary, the experiments suggest that a large part of the $\mathcal{R}$ spread across today's GSRMs is explained by different formulations of small-scale

mixing and cloud microphysical processes rather than by differences in model resolution. I could not test to what extent differences in the dynamical core, in particular the methods used to solve the governing equations, systematically influence $\mathcal{R}$. However, the fact that the sensitivity experiments produce a similar spread as the inter-model spread in the DYAMOND models (which do differ in their dynamical cores) as well as the modest influence of model resolution suggest that this influence is not large.

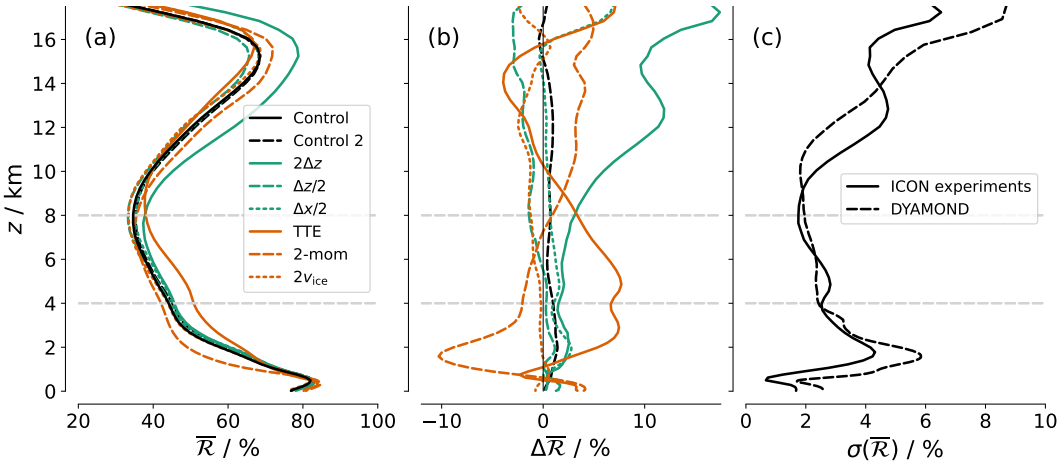


Figure 7: Tropical mean relative humidity $\overline{\mathcal{R}}$ in the control and sensitivity experiments shown in (a) absolute units and (b) as difference to the control experiment. (c) Standard deviation of $\mathcal{R}$ across all experiments (solid) and across the DYAMOND multi-model ensemble (dashed).

5.2 PHYSICAL MECHANISMS CONTROLLING RELATIVE HUMIDITY CHANGES

To better understand the physical mechanisms behind the $\mathcal{R}$ changes, I perform two types of Lagrangian reconstructions of $\mathcal{R}$ based on back-trajectories. The first reconstruction is an implementation of the last-saturation paradigm (Section 2.3). The part of the $\mathcal{R}$ differences that is explained by this reconstruction is caused by differences in last-saturation and/or target temperature (Equation 5). One could also argue that this part is caused by differences in the resolved (gridscale) circulation and temperature structure, which determine where saturation occurs. However, I will show that this is not always the case, because parameterized (sub-gridscale) processes like turbulent mixing can also directly influence last-saturation statistics. The second reconstruction additionally accounts for processes neglected in the last-saturation paradigm, i.e. any moisture sources and sinks that affect the humidity of air parcels after last-saturation. These include moistening from evaporation or sublimation of hydrometeors, as well as moistening or drying from sub-gridscale turbulent mixing, both of which are parameterized. Based on the two Lagrangian reconstructions, the actual ICON-simulated $\mathcal{R}$ changes can be decomposed into contributions from a) changes in last-saturation and/or target temperatures, b) changes in the parameterized moistening or drying after

last-saturation and c) changes in any processes not captured by either Lagrangian reconstruction. The latter include changes in the strength of numerical diffusion in the ICON model and imperfections in the calculated trajectories, e.g. due to limited temporal and spatial resolution of the model output fields the trajectories are calculated on. Since the last-saturation paradigm has never been used before to examine humidity differences between models, nor has it been implemented based on wind fields of storm-resolving resolution, this work also tests the quality of the paradigm in explaining differences between models as they begin to resolve the spectrum of vertical motions in the atmosphere.

For this part of the study I focus on the mid troposphere, more precisely the altitude layer between 4 km and 8 km, where I showed $\mathcal{R}$ differences to be of particular relevance in my first study. For each experiment (except from the $\Delta x/2$ experiment, due to limited resources), a large ensemble of back-trajectories is released once per day from random positions in the tropical mid troposphere. They are calculated backward for 15 days based on hourly model wind fields, which are re-gridded from the native ICON grid to a regular $0.1° \times 0.1°$ latitude-longitude grid. For each trajectory, points of last-saturation as well as subsequent moisture sources and sinks are determined based on moisture tendencies produced by the microphysics and turbulence schemes. A comparison shows that the $\mathcal{R}$ reconstruction including moisture sources and sinks captures the ICON-simulated $\mathcal{R}$ distribution in the control experiment slightly better than the plain last-saturation reconstruction, as is expected if the approach worked as intended.

To distinguish between different tropical humidity regimes, the mid troposphere is depicted in "$\mathcal{R}$-Space", i.e. it is split into ten equal-sized bins of $\mathcal{R}$. This representation reveals that $\mathcal{R}$ differences across experiments are generally larger in moist than in dry regions (Figure 8a) – similar to the $\mathcal{R}$ differences across the DYAMOND ensemble (Figure 6a). The depiction in $\mathcal{R}$-Space also shows that internal variability in $\mathcal{R}$, which is measured by the difference between the control experiment and the second control experiment with perturbed initial conditions (Control 2), increases towards dry regions, where it is of similar magnitude than the $\mathcal{R}$ differences in most sensitivity experiments. Thus, the $\mathcal{R}$ differences in the dry regions are strongly coloured by internal variability and longer experiments would be needed to quantify systematic differences resulting from differences in model formulation.

For most experiments the reconstruction based on the plain last-saturation model captures the actual, ICON-simulated $\mathcal{R}$ change compared to the control simulation to a large extent (Figure 8b), whereas changes in parameterized moisture sources and sinks only explain a minor part of the $\mathcal{R}$ change (Figure 8c). This means that $\mathcal{R}$ differences between experiments are mainly caused by variations in source and/or target temperature. A further decomposition shows that the contribution from changes in source temperature to the $\mathcal{R}$ change is generally larger than that from changes in target temperature (not shown here). These results highlight the importance of source temperatures compared to subsequent

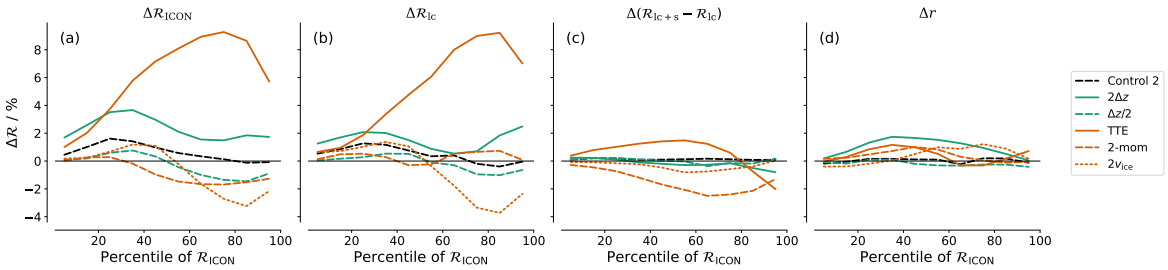


Figure 8: ICON-simulated and reconstructed mid-tropospheric $\mathcal{R}$ differences between experiments displayed in $\mathcal{R}$-Space. (a) ICON-simulated $\mathcal{R}$ differences to the control experiment. (b) $\mathcal{R}$ differences reconstructed from the plain last-saturation model and (c) differences in the effect from moisture sources and sinks after last-saturation. (d) Differences in the residual, i.e. in the difference between ICON-simulated and reconstructed $\mathcal{R}$. The sum of the terms shown in (b) to (d) yields the ICON-simulated $\mathcal{R}$ differences shown in (a).

moisture sources or sinks in causing $\mathcal{R}$ differences. In the particular case of the TTE turbulence experiment, this means that the strong mid-tropospheric moistening is not a direct consequence of enhanced vertical turbulent mixing that moistens air parcels as they are transported from source to target regions. Instead, it must be explained by changes in the properties of the source regions themselves.

There are two exceptions where the plain last-saturation model does not explain a major part of the $\mathcal{R}$ change. Only a part of the moistening in the experiment with coarser vertical resolution is captured by the plain last-saturation reconstruction (Figure 8b) and the remaining part is not explained by changes in parameterized moisture sources or sinks, as apparent from the rather large residuum term (Figure 8c). This could indicate that some of the moistening results from an increase in numerical diffusion associated with the decrease in vertical resolution. Also, the (rather weak) mid-tropospheric drying in the 2-mom experiment is mainly due to a reduction in evaporation of cloud condensate and/or precipitation after last-saturation. However, additional trajectory calculations show that the stronger lower-tropospheric drying in the 2-mom experiment (cf. Figure 7) is largely explained by the last-saturation model. The local moistening from shallow convection in this layer is weaker in the 2-mom experiment, which is consistent with the microphysics limiting the depth of shallow convection mentioned above.

As changes in last-saturation temperature turn out to be the main driver of $\mathcal{R}$ differences across experiments, I further examine their origin by separating the contributions from changes in inner-tropical and extra-tropical last-saturation temperatures. This reveals that the $\mathcal{R}$ changes in moist regions (roughly the moistest 60 % in $\mathcal{R}$ space) are dominated by changes in inner-tropical last-saturation temperatures. For the driest 40 % of the tropics, however, changes in tropical and extra-tropical source temperatures are of similar relevance. In the rather short experiments considered in this study, changes in the source temperature of dry regions are explained by, or can at least not be distinguished from, internal variability. This variability likely reflects variations in the large-scale circulation that

connects the dry target regions to their remote source regions. Internal variability is particularly strong in extra-tropical source temperatures. This highlights the importance of transport pathways by which extra-tropical air is brought into the tropics in controlling the humidity of the dry regions (see Aemisegger et al., 2021; Cau et al., 2007; Roca et al., 2012).

To better understand the strong impact of the turbulence parameterization on last-saturation temperature, I compare moisture tendencies produced by the turbulence and microphysics schemes in the control and TTE experiments. The comparison reveals that, particularly in moist tropical regimes, the TTE scheme performs a much more vigorous vertical mixing between boundary layer and free troposphere than the Smagorinsky-Lilly scheme. This results in a strong turbulent moistening of the mid troposphere, which favours condensation over broader areas than in the control experiment. In other words, the TTE scheme creates a cloudier inner-tropical mid-troposphere, which changes last-saturation statistics such that the share of warm last-saturation temperatures increases. This demonstrates that last-saturation statistics are not determined exclusively by the resolved circulation and temperature structure, but can also be directly affected by parameterized processes.

By exposing the strong mixing between boundary layer and free troposphere performed by the TTE turbulence scheme, my analysis makes an important contribution to the model development at our institute. There is an ongoing discussion on which type of turbulence scheme to use in storm-resolving simulations. While the TTE scheme has been used in the GCM version of ICON, the Smagorinsky-Lilly scheme is designed for large-eddy simulations. My analysis reveals that the TTE scheme, in its current form, behaves similar to a convective parameterization when employed at storm-resolving scales, which was not intended. This highlights that a better adaption of existing turbulence schemes is needed if they are to be used for storm-resolving simulations. However, the fact that even the extreme perturbation applied by switching between different types of turbulence schemes did not change $\mathcal{R}$ far beyond the inter-model spread in DYAMOND is promising. Many of the DYAMOND models used turbulence parameterizations that were not specifically adapted to storm-resolving scales due to their early development stage. Hence, a better adaption of the schemes in future model versions might further reduce the spread in tropical $\mathcal{R}$.

In summary, I find that mid-tropospheric $\mathcal{R}$ changes in the sensitivity experiments, including the strong moistening in the turbulence experiment, are largely explained by changes in last-saturation temperature. In the moist regions of the tropics $\mathcal{R}$ is only sensitive to changes in inner-tropical last-saturation statistics and hence depends on the representation of the tropical deep convective regions. The latter is strongly influenced by the parameterization of small-scale turbulent mixing. In the dry regions $\mathcal{R}$ is equally sensitive to changes in inner-tropical and extra-tropical last-saturation statistics and depends on the large-scale circulation, in particular on the mechanisms of exchange between the tropics and extra-tropics.

The experiments are too short to say with certainty which model uncertainties have the greatest impact on these mechanisms.

In a nutshell, the second study yields the following key results:

- Model parameterizations, rather than model resolution, represent the major source of relative humidity spread across global storm-resolving models.
- Vertical mixing processes strongly impact the relative humidity of the moist tropics by changing last-saturation statistics within the tropics rather than by changing the evolution of humidity after last-saturation.
- The humidity of the dry tropics is disproportionately sensitive to changes in the pathways of exchange with the extra-tropics.

6

SUMMARY, CONCLUSIONS AND OUTLOOK

The studies presented in this essay are the first to examine uncertainties in modelling the distribution of relative humidity in the tropics at kilometer-scale resolution. In the first study, I use a first intercomparison of GSRMs (DYAMOND) to quantify differences in relative humidity and assess their relevance by calculating their effect on clear-sky OLR. In the second study, I narrow down the sources of these differences by examining how much and through which physical mechanisms relative humidity is affected by changes in model resolution and parameterizations.

The first study revealed that the inter-model spread in tropical mean relative humidity is reduced (approximately halved in the free troposphere) compared to classical GCMs with convective parameterizations. This is a promising result, particularly in light of the fact that the GSRMs that participated in DYAMOND were not specifically tuned for that experiment. Considering the results of the second study that highlighted the importance of variations in last-saturation (or "source") temperatures in causing humidity differences, this could indicate that the source regions of tropical air, indeed the convective regions, are better constrained in GSRMs than in GCMs with convective parameterizations.

To quantify the uncertainty of relative humidity in GSRMs, besides comparing the models to each other, they also need to be evaluated with observations. A first comparison to the ERA5 reanalysis in my first study indicates that GSRMs as a group are at least not systematically biased with respect to observations. Nevertheless, to improve individual models their biases need to be known precisely. Since the humidity in (re-)analyses can be affected by errors in the underlying forecast model (e.g. Pincus et al., 2017) and ERA5 is known to exhibit a moist bias with respect to satellite observations (Xue et al., 2020), further comparisons should also be made to purely observational data sets, e.g. from radiosondes and from satellite instruments such as the humidity sounders I briefly introduced in this thesis.

My analysis is limited to differences in the present-day humidity distribution and their effect on the present-day radiation budget. A question that I cannot answer yet is whether the inter-model spread in the tropical clear-sky feedback is also reduced across GSRMs. However, there are reasons to be optimistic that the feedback is more consistent across GSRMs. On the one hand, to the extent that the feedback directly depends on the present-day humidity (Bourdin et al., 2021; McKim et al., 2021), the reduced spread implies a more consistent feedback. On the other hand, the feedback depends on the change of relative humidity under warming. It seems unlikely that the spread in the relative humidity response

is increased although the present-day relative humidity is better constrained in GSRMs. This needs to be verified as soon as model intercomparisons are performed for various climate states ("time slice experiments") or even over climatological time periods.

Although the model spread in relative humidity is reduced across GSRMs, it still causes a considerable spread (a standard deviation of $\sim 1.2\,\mathrm{W\,m^{-2}}$) in tropical mean clear-sky OLR. I identified the lower and mid free troposphere as the regions where future reductions of the spread in relative humidity would be most beneficial to reduce uncertainty in clear-sky OLR. This applies both to the comparably small relative humidity differences in dry regimes and the large differences in moist regimes. However, eventually the clear-sky effect is expected to be more relevant in the dry regimes, where upper-level clouds with a potential masking effect are rare.

The sensitivity experiments in my second study suggest that the remaining model spread in lower- and mid-tropospheric relative humidity is mainly driven by differences in parameterizations rather than differences in model resolution. In the limited number of sensitivity experiments, changes in the microphysics parameterization most strongly affect relative humidity in the lower-tropospheric region associated with shallow clouds, whereas changing the turbulence parameterization changes relative humidity in a broad layer in the mid troposphere. The humidity differences produced in the sensitivity experiments, including the most pronounced ones caused by parameterization changes, are to a large extent explained by differences in last-saturation temperature, whereas differences in moisture sources or sinks after last-saturation play a minor role. This highlights the importance of correctly capturing the source temperature of tropical mid-tropospheric air, that is both the temperature structure and the circulation.

For the moist regimes of the tropical mid-troposphere (roughly the moist half in terms of mid-tropospheric relative humidity) last-saturation takes place in nearby mid-level or deep clouds. The sensitivity experiments indicate that a major uncertainty in modelling the area covered by such clouds lies in the representation of small-scale turbulent mixing. Today's GSRMs use turbulence parameterizations that were either designed for higher-resolution large-eddy simulations or coarser-resolution GCMs, but neither of them have been specifically adapted to simulations at kilometer-scales. A better adaption of the schemes is therefore an important step towards reducing the uncertainty in the representation of the convective source regions.

For the dry half of the tropical mid troposphere, source regions are located more remotely, both in the upper-level outflow regions of inner-tropical deep convection and in the extra-tropical upper troposphere. Source temperatures are therefore influenced by large-scale circulation patterns, which vary on longer time scales than the simulations I could conduct. Hence, while for the moist tropics one simulated month is sufficient to distinguish systematic changes in source

temperature from internal variability, longer simulations are needed to understand the sources of spread in the dry tropics. This represents an important limitation of my study. At the same time it highlights the importance of different large-scale transport pathways in controlling relative humidity in the dry tropical regions, particularly the pathways from the extra-tropics into the tropics (see also Aemisegger et al., 2021; Cau et al., 2005; Roca et al., 2012). Understanding these transport pathways, how they might change under warming and how they differ across models is therefore key to reduce the uncertainty in the clear-sky climate feedback.

An important step to reduce the uncertainty in the source temperature of tropical free-tropospheric air is to constrain the characteristics of the main source regions, i.e. the deep convective regions, with observations. Satellite, airborne and ground-based observations can provide constraints on quantities related to last-saturation statistics, such as the depth and amount of convective clouds, the height and extent of anvil clouds as well as on the circulation and thermal structure of the atmosphere. Apart from existing satellite missions and measurement sites, valuable data will be provided by new satellite missions such as the Earth Cloud Aerosol and Radiation Explorer (EarthCARE) mission (Illingworth et al., 2015), as well as upcoming measurement campaigns focusing on the ITCZ region, such as the "EarthCARE, Tropical Oceans and Organized Convection" (EC-TOOC) campaign planned for the year 2024 in the tropical Atlantic. A technique to infer the meso-scale circulation from dropsonde measurements distributed along a circular flight path was developed for the EUREC^{4}A campaign. Applying this technique to deep convective regimes could allow to determine the altitude at which air detrains from deep convection very precisely.

Besides the structure of the source regions themselves, the transport pathways by which air is brought to the tropical free troposphere need to be evaluated. This could be done using water vapor isotopologues. Since the isotopic composition of water vapor is affected by phase changes of water, measurements of this composition can provide information on the origin of air, in particular on pathways of transport from the extra-tropics into the tropics (e.g. Aemisegger et al., 2021; González et al., 2016), which affect the humidity of the dry zones. Incorporating isotopes in models and comparing modelled to observed isotopic signatures, as performed by Risi et al., 2012 for GCMs, could be a promising approach to identify differences between modelled and observed pathways.

In summary, already in their early development stage GSRMs show better agreement in tropical free-tropospheric humidity than classical GCMs with convective parameterizations, and hence represent a promising tool to reduce the uncertainty in the tropical clear-sky climate feedback. To exploit their full potential, more research is needed on adapting the remaining parameterizations for use at kilometer-scale resolutions. A particular focus should be placed on a realistic representation of the source regions and the transport pathways of air arriving in the tropical mid troposphere.

BIBLIOGRAPHY

Aemisegger, Franziska et al. (2021). "How Rossby wave breaking modulates the water cycle in the North Atlantic trade wind region." In: *Weather and Climate Dynamics* 2.1, pp. 281–309. DOI: 10.5194/wcd-2-281-2021.

Albright, Anna Lea et al. (2022). "Observed Subcloud-Layer Moisture and Heat Budgets in the Trades." In: *Journal of the Atmospheric Sciences* 79.9, pp. 2363–2385. DOI: 10.1175/JAS-D-21-0337.1.

Bony, S. et al. (2006). "How Well Do We Understand and Evaluate Climate Change Feedback Processes?" In: *Journal of Climate* 19.15, pp. 3445–3482. DOI: 10.1175/JCLI3819.1.

Bony, Sandrine (2005). "Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models." In: *Geophysical Research Letters* 32.20. DOI: 10.1029/2005GL023851.

Bourdin, S., L. Kluft, and B. Stevens (2021). "Dependence of Climate Sensitivity on the Given Distribution of Relative Humidity." In: *Geophysical Research Letters*. DOI: https://doi.org/10.1029/2021GL092462.

Bretherton, C. S., P. N. Blossey, and M. Khairoutdinov (2005). "An Energy-Balance Analysis of Deep Convective Self-Aggregation above Uniform SST." In: *Journal of the Atmospheric Sciences* 62.12, pp. 4273–4292. DOI: 10.1175/JAS3614.1.

Buehler, S. A. (2005). "A simple method to relate microwave radiances to upper tropospheric humidity." In: *Journal of Geophysical Research* 110.D2. DOI: 10.1029/2004JD005111.

Buehler, S. A. et al. (2008). "An upper tropospheric humidity data set from operational satellite microwave data." In: *Journal of Geophysical Research* 113.D14. DOI: 10.1029/2007JD009314.

Cau, Piero, John Methven, and Brian Hoskins (2005). "Representation of dry tropical layers and their origins in ERA-40 data." In: *Journal of Geophysical Research: Atmospheres* 110.D6, n/a–n/a. DOI: 10.1029/2004JD004928.

— (2007). "Origins of Dry Air in the Tropics and Subtropics." In: *Journal of Climate* 20.12, pp. 2745–2759. DOI: 10.1175/JCLI4176.1.

Colman, Robert and Brian J. Soden (2021). "Water vapor and lapse rate feedbacks in the climate system." In: *Reviews of Modern Physics* 93.4. DOI: 10.1103/RevModPhys.93.045002.

Dessler, A. E. and S. C. Sherwood (2000). "Simulations of tropical upper tropospheric humidity." In: *Journal of Geophysical Research* 105.D15, pp. 20155–20163. DOI: 10.1029/2000JD900231.

Emanuel, K. A. and R. T. Pierrehumbert (1996). "Clouds, Chemistry and Climate." In: ed. by P.J. Crutzen and V. Ramanathan. Springer-Verlag Berlin Heidelberg. Chap. Chapter 2: Microphysical and Dynamical Control of Tropospheric Water Vapor, pp. 17–28.

Galewsky, J., A. Sobel, and I. Held (2005). "Diagnosis of Subtropical Humidity Dynamics Using Tracers of Last Saturation." In: *Journal of the Atmospheric Sciences* 62.9, pp. 3353–3367. DOI: 10.1175/JAS3533.1.

George, Geet et al. (2021). "JOANNE: Joint dropsonde Observations of the Atmosphere in tropical North atlaNtic meso-scale Environments." In: *Earth System Science Data* 13.11, pp. 5253–5272. DOI: 10.5194/essd-13-5253-2021.

González, Yenny et al. (2016). "Detecting moisture transport pathways to the subtropical North Atlantic free troposphere using paired H&ampltsub&ampgt2&amplt/sub&ampgtO-&amplti&ampgt&amplt/i&ampgtD in situ measurements." In: *Atmospheric Chemistry and Physics* 16.7, pp. 4251–4269. DOI: 10.5194/acp-16-4251-2016.

Goody, R. M. and G. D. Robinson (1951). "Radiation in the troposphere and lower stratosphere." In: *Quarterly Journal of the Royal Meteorological Society* 77.332, pp. 151–187. DOI: 10.1002/qj.49707733203.

Held, I. M. and B. J. Soden (2000). "Water Vapour Feedback and Global Warming." In: *Annual Review of Energy and the Environment* 25.1, pp. 441–475. DOI: 10.1146/annurev.energy.25.1.441.

Hersbach, Hans et al. (2020). "The ERA5 global reanalysis." In: *Quarterly Journal of the Royal Meteorological Society* 146.730, pp. 1999–2049. DOI: 10.1002/qj.3803.

Hohenegger, Cathy et al. (2022). "ICON-Sapphire: simulating the components of the Earth System and their interactions at kilometer and subkilometer scales." In: DOI: 10.5194/gmd-2022-171.

Huang, Yi and Maziar Bani Shahabadi (2014). "Why logarithmic? A note on the dependence of radiative forcing on gas concentration." In: *Journal of Geophysical Research: Atmospheres* 119.24, pp. 13,683–13,689. DOI: 10.1002/2014JD022466.

IPCC (2021). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by V. Masson-Delmotte et al. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. DOI: 10.1017/9781009157896.

Illingworth, A. J. et al. (2015). "The EarthCARE Satellite: The Next Step Forward in Global Measurements of Clouds, Aerosols, Precipitation, and Radiation." In: *Bulletin of the American Meteorological Society* 96.8, pp. 1311–1332. DOI: 10.1175/BAMS-D-12-00227.1.

Ingram, William (2010). "A very simple model for the water vapour feedback on climate change." In: *Quarterly Journal of the Royal Meteorological Society* 136.646, pp. 30–40. DOI: 10.1002/qj.546.

Jeevanjee, Nadir, Daniel D. B. Koll, and Nicholas Lutsko (2021). ""Simpson's Law" and the Spectral Cancellation of Climate Feedbacks." In: *Geophysical Research Letters* 48.14. DOI: 10.1029/2021GL093699.

John, V. O. and B. J. Soden (2007). "Temperature and humidity biases in global climate models and their impact on climate feedbacks." In: *Geophysical Research Letters* 34.18, p. L18704. DOI: 10.1029/2007GL030429.

Kiehl, J. T. and Kevin E. Trenberth (1997). "Earth's Annual Global Mean Energy Budget." In: *Bulletin of the American Meteorological Society* 78.2, pp. 197–208. DOI: 10.1175/1520-0477(1997)078<0197:EAGMEB>2.0.CO;2.

Kluft, Lukas et al. (2019). "Re-Examining the First Climate Models: Climate Sensitivity of a Modern Radiative–Convective Equilibrium Model." In: *Journal of Climate* 32.23, pp. 8111–8125. DOI: 10.1175/JCLI-D-18-0774.1.

Koll, Daniel D. B. and Timothy W. Cronin (2018). "Earth's outgoing longwave radiation linear due to H2O greenhouse effect." In: *Proceedings of the National Academy of Sciences* 115.41, pp. 10293–10298. DOI: 10.1073/pnas.1809868115.

Krautstrunk, Monika and Andreas Giez (2012). "The Transition From FALCON to HALO Era Airborne Atmospheric Research." In: pp. 609–624. DOI: 10.1007/978-3-642-30183-4_37.

Mauritsen, Thorsten et al. (2012). "Tuning the climate of a global model." In: *Journal of Advances in Modeling Earth Systems* 4.3, n/a–n/a. DOI: 10.1029/2012MS000154.

McKim, Brett A., Nadir Jeevanjee, and Geoffrey K. Vallis (2021). "Joint Dependence of Longwave Feedback on Surface Temperature and Relative Humidity." In: *Geophysical Research Letters* 48.18. DOI: 10.1029/2021GL094074.

Medeiros, Brian, Bjorn Stevens, and Sandrine Bony (2014). "Using aquaplanets to understand the robust responses of comprehensive climate models to forcing." In: *Climate Dynamics* 44.7-8, pp. 1957–1977. DOI: 10.1007/s00382-014-2138-0.

Nakajima, Shinichi, Yoshi-Yuki Hayashi, and Yutaka Abe (1992). "A Study on the "Runaway Greenhouse Effect" with a One-Dimensional Radiative–Convective Equilibrium Model." In: *Journal of the Atmospheric Sciences* 49.23, pp. 2256–2266. DOI: 10.1175/1520-0469(1992)049<2256:ASOTGE>2.0.CO;2.

O'Gorman, P. A. and C. J. Muller (2010). "How closely do changes in surface and column water vapor follow Clausius–Clapeyron scaling in climate change simulations?" In: *Environmental Research Letters* 5.2, p. 025207. DOI: 10.1088/1748-9326/5/2/025207.

Petty, G. W. (2006). *A First Course in Atmospheric Radiation*. Sundog Pub. ISBN: 0972903313.

Pierce, D. W. et al. (2006). "Three-dimensional tropospheric water vapor in coupled climatemodels compared with observations from the AIRS satellite system." In: *Geophysical Research Letters* 33.21, p. L21701. DOI: 10.1029/2006GL027060.

Pierrehumbert, R. T. (1995). "Thermostats, Radiator Fins, and the Local Runaway Greenhouse." In: *Journal of the Atmospheric Sciences* 52.10, pp. 1784–1806. DOI: 10.1175/1520-0469(1995)052<1784:TRFATL>2.0.CO;2.

Pierrehumbert, R. T., H. Brogniez, and R. Roca (2007). "The Global Circulation of the Atmosphere." In: ed. by T. Schneider and A. Sobel. Princeton University Press. Chap. On the relative humidity of the atmosphere, pp. 143–185.

Pierrehumbert, R. T. and R. Roca (1998). "Evidence for Control of Atlantic Subtropical Humidity by Large Scale Advection." In: *Geophysical Research Letters* 25.24, pp. 4537–4540. DOI: 10.1029/1998GL900203.

Pierrehumbert, Raymond T. (1999). "Subtropical water vapor as a mediator of rapid global climate change." In: pp. 339–361. DOI: 10.1029/GM112p0339.

Pincus, R. et al. (2017). "The Representation of Tropospheric Water Vapor Over Low-Latitude Oceans in (Re-)analysis: Errors, Impacts, and the Ability to Exploit Current and Prospective Observations." In: *Surveys in Geophysics* 38.6, pp. 1399–1423. DOI: 10.1007/s10712-017-9437-z.

Po-Chedley, Stephen et al. (2018). "Sources of Intermodel Spread in the Lapse Rate and Water Vapor Feedbacks." In: *Journal of Climate* 31.8, pp. 3187–3206. DOI: 10.1175/JCLI-D-17-0674.1.

Po-Chedley, S. et al. (2019). "Climatology Explains Intermodel Spread in Tropical Upper Tropospheric Cloud and Relative Humidity Response to Greenhouse Warming." In: *Geophysical Research Letters* 46.22, pp. 13399–13409. DOI: 10.1029/2019GL084786.

Risi, Camille et al. (2012). "Process-evaluation of tropospheric humidity simulated by general circulation models using water vapor isotopic observations: 2. Using isotopic diagnostics to understand the mid and upper tropospheric moist bias in the tropics and subtropics." In: *Journal of Geophysical Research: Atmospheres* 117.D5. DOI: 10.1029/2011JD016623.

Roca, R. et al. (2012). "Tropical and Extra-Tropical Influences on the Distribution of Free Tropospheric Humidity Over the Intertropical Belt." In: *Surveys in Geophysics* 33, pp. 565–583. DOI: 10.1007/s10712-011-9169-4.

Romps, D. M. (2014). "An Analytical Model for Tropical Relative Humidity." In: *Journal of Climate* 27.19, pp. 7432–7449. DOI: 10.1175/JCLI-D-14-00255.1.

Satoh, M. et al. (2019). "Global Cloud-Resolving Models." In: *Current Climate Change Reports* 5.3, pp. 172–184. DOI: 10.1007/s40641-019-00131-0.

Schulz, H. and B. Stevens (2018). "Observing the Tropical Atmosphere in Moisture Space." In: *Journal of the Atmospheric Sciences* 75.10, pp. 3313–3330. DOI: 10.1175/JAS-D-17-0375.1.

Seeley, Jacob T. and Nadir Jeevanjee (2021). "H sub2/sub O Windows and CO sub2/sub Radiator Fins: A Clear-Sky Explanation for the Peak in Equilibrium Climate Sensitivity." In: *Geophysical Research Letters* 48.4. DOI: 10.1029/2020GL089609.

Sherwood, S. C. (1996). "Maintainance of the Free-Tropospheric Tropical Water Vapor Distribution. Part II: Simulation by Large-Scale Advection." In: *Journal of Climate* 9.11, pp. 2919–2934. DOI: 10.1175/1520-0442(1996)009<2919:MOTFTT>2.0.CO;2.

Sherwood, S. C. et al. (2010). "Relative humidity changes in a warmer climate." In: *Journal of Geophysical Research* 115.D9, p. D09104. DOI: 10.1029/2009JD012585.

Sherwood, Steven C., Sandrine Bony, and Jean-Louis Dufresne (2014). "Spread in model climate sensitivity traced to atmospheric convective mixing." In: *Nature* 505.7481, pp. 37–42. DOI: 10.1038/nature12829.

Shi, Lei and John J. Bates (2011). "Three decades of intersatellite-calibrated High-Resolution Infrared Radiation Sounder upper tropospheric water vapor." In: *Journal of Geophysical Research* 116.D4. DOI: 10.1029/2010JD014847.

Shine, Keith P. and Ashok Sinha (1991). "Sensitivity of the Earth's climate to height-dependent changes in the water vapour mixing ratio." In: *Nature* 354.6352, pp. 382–384. DOI: 10.1038/354382a0.

Simpson, G. C. (1928). "Some studies in terrestrial radiation." In: *Memoirs of the Royal Meteorological Society* 2.16, pp. 69–95.

Soden, B. J. and I. M. Held (2006). "An Assessment of Climate Feedbacks in Coupled Ocean–Atmosphere Models." In: *Journal of Climate* 19.14, pp. 3354–3360. DOI: 10.1175/JCLI3799.1.

Soden, Brian J. and Francis P. Bretherton (1993). "Upper tropospheric relative humidity from the GOES 6.7 μm channel: Method and climatology for July 1987." In: *Journal of Geophysical Research* 98.D9, p. 16669. DOI: 10.1029/93JD01283.

Spencer, R. W. and W. D. Braswell (1997). "How Dry is the Tropical Free Troposphere? Implications for Global Warming Theory." In: *Bulletin of the American Meteorological Society* 78.6, pp. 1097–1106. DOI: 10.1175/1520-0477(1997)078<1097:hdittf>2.0.co;2.

Stevens, B. and S. Bony (2013a). "What Are Climate Models Missing?" In: *Science* 340.6136, pp. 1053–1054. DOI: 10.1126/science.1237554.

Stevens, Bjorn and Sandrine Bony (2013b). "Water in the atmosphere." In: *Physics Today* 66.6, pp. 29–34. DOI: 10.1063/PT.3.2009.

Stevens, Bjorn et al. (2019). "DYAMOND: the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains." In: *Progress in Earth and Planetary Science* 6.61. DOI: 10.1186/s40645-019-0304-z.

Stevens, Björn and Axel Seifert (2008). "Understanding macrophysical outcomes of microphysical choices in simulations of shallow cumulus convection." In: *Journal of the Meteorological Society of Japan* 86A, pp. 143–162. DOI: 10.2151/jmsj.86A.143.

Taylor, Karl E., Ronald J. Stouffer, and Gerald A. Meehl (2012). "An Overview of CMIP5 and the Experiment Design." In: *Bulletin of the American Meteorological Society* 93.4, pp. 485–498. DOI: 10.1175/BAMS-D-11-00094.1.

Tomita, H. and M. Satoh (2004). "A new dynamical framework of nonhydrostatic global model using the icosahedral grid." In: *Fluid Dynamics Research* 34.6, pp. 357–400. DOI: 10.1016/j.fluiddyn.2004.03.003.

Tsushima, Yoko et al. (2020). "Investigating physical constraints on climate feedbacks using a perturbed parameter ensemble." In: *Climate Dynamics* 55.5-6, pp. 1159–1185. DOI: 10.1007/s00382-020-05318-y.

Vial, J., J.-L. Dufresne, and S. Bony (2013). "On the interpretation of inter-model spread in CMIP5 climate sensitivity estimates." In: *Climate Dynamics* 41.11-12, pp. 3339–3362. DOI: 10.1007/s00382-013-1725-9.

Vial, Jessica et al. (2016). "Coupling between lower-tropospheric convective mixing and low-level clouds: Physical mechanisms and dependence on convection scheme." In: *Journal of Advances in Modeling Earth Systems* 8.4, pp. 1892–1911. DOI: 10.1002/2016MS000740.

Xue, Y. et al. (2020). "Assessment of Upper Tropospheric Water Vapor Monthly Variation in Reanalyses With Near-Global Homogenized 6.5-μm Radiances From Geostationary Satellites." In: *Journal of Geophysical Research: Atmospheres* 125.18. DOI: 10.1029/2020JD032695.

Zängl, Günther et al. (2015). "The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core." In: *Quarterly Journal of the Royal Meteorological Society* 141.687, pp. 563–579. DOI: 10.1002/qj.2378.

Part II

APPENDIX

A

TROPICAL FREE-TROPOSPHERIC HUMIDITY DIFFERENCES AND THEIR EFFECT ON THE CLEAR-SKY RADIATION BUDGET IN GLOBAL STORM-RESOLVING MODELS

The work in this appendix has been published as:

Lang, T., Naumann, A. K., Stevens, B., and Buehler, S. A. (2021). Tropical free-tropospheric humidity differences and their effect on the clear-sky radiation budget in global storm-resolving models. *Journal of Advances in Modeling Earth Systems*, 13, e2021MS002514. https://doi.org/10.1029/2021MS002514

CONTRIBUTIONS Together with all co-authors I conceptualized the study. A. K. Naumann and S. A. Buehler supervised the study. I conducted the analysis and prepared the manuscript with contributions from all co-authors.

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 Wm^{-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

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.

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.

A.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 and Soden, 2000; Po-Chedley et al., 2019; Romps, 2014). 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. Bony et al., 2006; Soden and Held, 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 (Po-Chedley et al., 2018; Vial et al., 2013). 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 and 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 and 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. Dessler and Sherwood, 2000; Pierrehumbert and Roca, 1998; Sherwood, 1996). 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 and 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 and Stevens, 2018) and to bypass the issue of co-location when comparing observations and model simulations (Naumann and Kiemle, 2020).

This paper is organized as follows: In Section A.2 we introduce the DYAMOND simulations and describe our post-processing of the model output. In Section A.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 A.4.

A.2 DYAMOND SIMULATIONS

A.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.

A.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 A.3.1). In the same manner we estimated the sampling uncertainty for each block in moisture space (Section A.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 A.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.

A.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.

A.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 A.9). 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 A.9c) 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 $\sigma(\mathrm{RH})$ is 2–3% RH and approximately constant with height. RH is lower than in ERA5 in the majority of models, except for ICON and NICAM. Above 10 km $\sigma(\mathrm{RH})$ increases with altitude and exceeds 8% RH at 100 hPa.

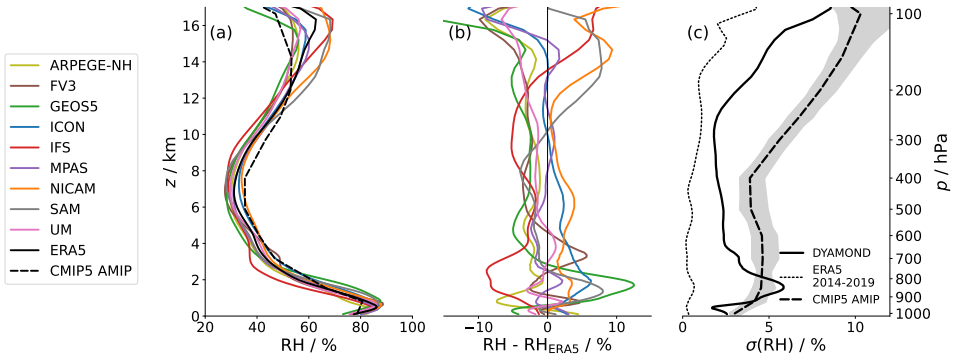


Figure A.9: Tropical mean RH profiles and inter-model spread in the DYAMOND ensemble. (a) Tropical mean vertical profiles of RH over ocean regions from all DYAMOND models (colours), the ERA5 reanalysis (black solid) and the CMIP5 AMIP 30-year multi-model mean (black dashed). (b) Vertical RH profiles 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.

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 A.10a,b). At lower levels, where T anomalies are small, q and RH anomalies are correlated (Figure A.9b, Figure A.10d). In the upper troposphere, where T anomalies are large, T and q anomalies are correlated (Figure A.10b,d), consistent with the idea that model errors in T cause errors in q. Although RH anomalies are also large there (Figure A.9), 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 A.9c. 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

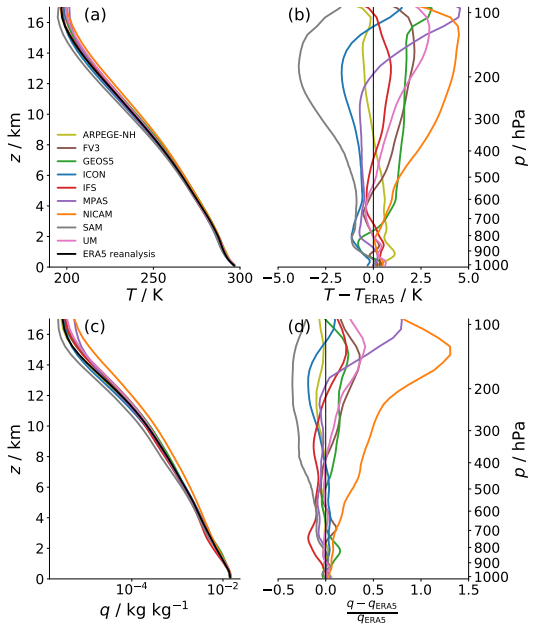


Figure A.10: Tropical mean vertical profiles of temperature T and specific humidity q over ocean regions from all DYAMOND models. Vertical profiles of T (a, b) and q (c, d) are shown as absolute values together with the ERA5 profiles (a, c) and as deviation from the ERA5 profiles (b, d). Deviations in q are in fractional units, i.e. normalized by the ERA5 value (q_{ERA5}).

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 A.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 for 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 A.11 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 A.9). 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 A.9c). 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 A.4 the RH differences in this region do not significantly affect the clear-sky radiation budget.

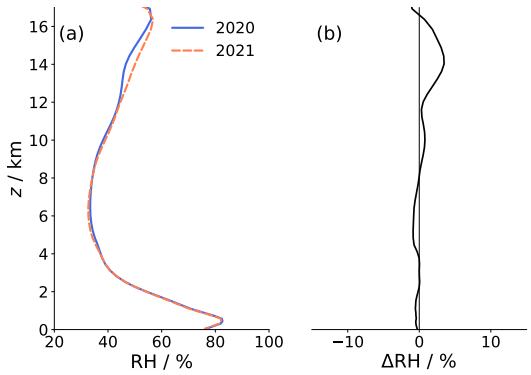


Figure A.11: Comparison of RH in two subsequent Februaries of a coupled atmosphere-ocean simulation with the ICON model at storm-resolving resolution (5 km). (a) Tropical mean (ocean only) RH in February 2020 (blue) and February 2021 (orange). February 2020 corresponds to days 13 to 40 after initialization, which is comparable to the analyzed DYAMOND period. (b) RH difference between February 2020 and February 2021.

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 A.9c. 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 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 $\sigma(\mathrm{RH})$ 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 A.9a). 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 A.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 A.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 A.4, the remaining RH differences still have a non-negligible impact on the clear-sky radiation budget.

A.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; Naumann and Kiemle, 2020; Schulz and Stevens, 2018). To span the moisture space, the randomly selected atmospheric profiles (Section A.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 A.12). 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 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. This block-averaging results in an x-axis that is linear in the percentile of IWV. Note

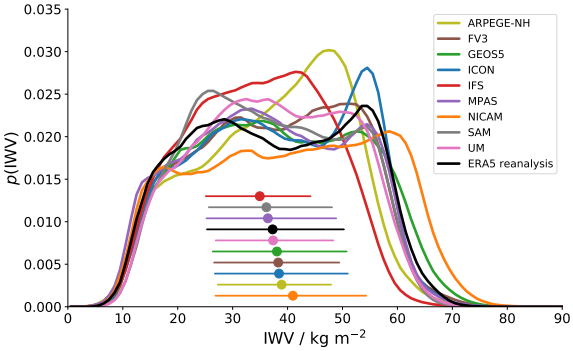


Figure A.12: Probability density function of integrated water vapor (IWV) over tropical ocean regions in the DYAMOND models and ERA5. Percentiles of each model's IWV distribution are shown below the curves: Coloured circles indicate the median, horizontal bars range from the 25th to the 75th percentile.

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. IWV itself increases rather linearly with the percentile (black line in Figure A.13d), but deviations in the upper- and lowermost percentiles are indicative of long tails in the IWV distribution (Figure A.12), and hence unusually potent moist and dry extremes.

SST increases from about 292 K in low IWV percentiles to about 302 K in high percentiles (Figure A.13d). 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 A.13c) 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 A.13c is truncated at 1.2 cm s^{-1}. 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)

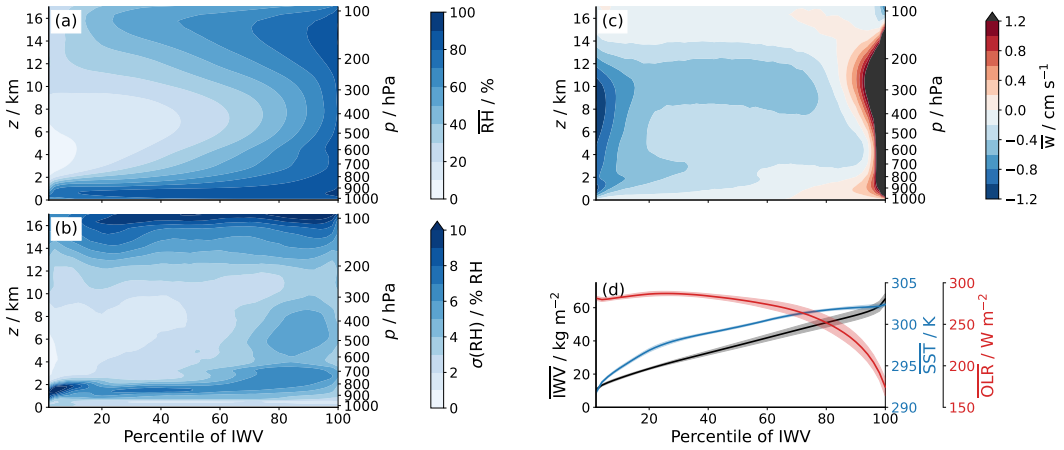


Figure A.13: Distributions of different block-averaged quantities in moisture space: (a) multi-model mean RH, (b) multi-model standard deviation of RH, (c) multi-model mean vertical velocity and (d) multi-model mean IWV (black), SST (blue) and all-sky OLR (red). Note that the color map for vertical velocity 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.

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 A.13d).

The largest RH values are found in the BL (Figure A.13a), 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, 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 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 A.16 in A.6.1. Here we focus on the inter-model standard deviation $\sigma(\mathrm{RH})$, shown in Figure A.13b. First, it is apparent that the large inter-model spread in the upper troposphere (Figure A.9) prevails throughout the entire tropics. In the tropopause region $\sigma(\mathrm{RH})$ exceeds 10% RH everywhere except from the driest part of the subsidence regions. Second, the local maximum in $\sigma(\mathrm{RH})$ 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 A.13a). In moister regions, where the RH gradient is less steep, the maximum in $\sigma(\mathrm{RH})$ is weaker but broader. Third, in the mid troposphere $\sigma(\mathrm{RH})$ 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 A.13c). 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.

A.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.

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 A.13d). 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.

A.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 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 A.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 and 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 A.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.

A.4.2 *Model differences in clear-sky OLR*

Tropical mean OLR_c differs by more than $4\,Wm^{-2}$ between the two most extreme models IFS and ICON (Figure A.14a). The multi-model standard deviation $\sigma(OLR_c)$ in tropical mean clear-sky OLR is $1.2\,Wm^{-2}$. 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 A.14b): One local maximum in $\sigma(OLR_c)$ 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 A.13b). A second, slightly weaker maximum in $\sigma(OLR_c)$ is located at the dry end of moisture space. In the next section we aim to better understand why the spread in OLR_c maximizes in these two regimes and which altitude regions in the troposphere contribute most.

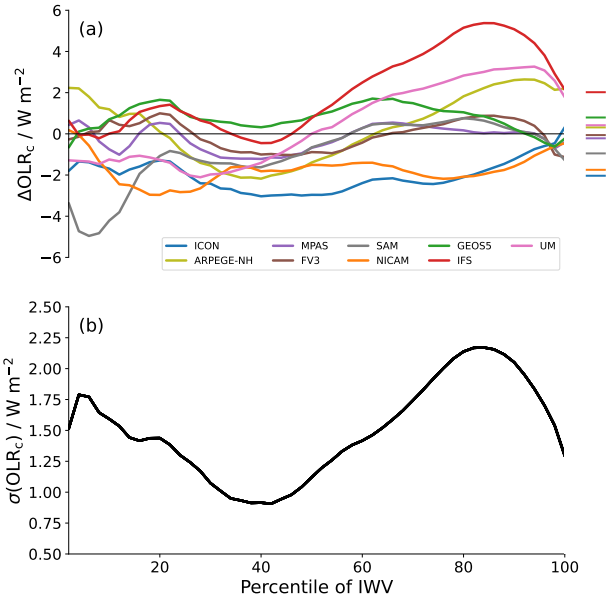


Figure A.14: Inter-model differences in clear-sky OLR in moisture space. (a) Anomalies in clear-sky OLR for each model, defined as the deviation from the ERA5 value and (b) inter-model standard deviation of clear-sky OLR.

A.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).

The atmosphere is divided into N vertical layers and it is assumed that a model's clear-sky OLR anomaly $\Delta\mathrm{OLR_c}$ can be expanded in a linear form as:

$$\Delta\mathrm{OLR_c} \approx \sum_{i=1}^{N} \left(K_i^e \, \Delta \mathrm{e}_i + K_i^T \, \Delta \mathrm{T}_i \right) \approx \sum_{i=1}^{N} K_i^{\mathrm{RH}} \, \Delta \mathrm{RH}_i, \tag{6}$$

where the index i denotes the vertical layer. K_i^x is the i-th component of the the vector $\mathbf{K^x}$, called radiative kernel. It describes the sensitivity of $\mathrm{OLR_c}$ to changes in a variable x in each layer i:

$$K_i^x = \frac{\partial \mathrm{OLR_c}}{\partial \mathrm{x}_i}. \tag{7}$$

The first approximation in Equation 6 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 A.17 in A.6.2. The computation of the kernels is also described in A.6.2.

Perturbations in e and T have opposite effects on $\mathrm{OLR_c}$, which is evident from the different signs of the respective kernels (Figure A.17). 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 (Ingram, 2010; Nakajima et al., 1992). 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 6.

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 $\mathbf{K^{RH,e}}$ and $\mathbf{K^{RH,T}}$, respectively:

$$
\begin{aligned}
K_i^{\mathrm{RH},e} &= \frac{\partial \mathrm{OLR_c}}{\partial \mathrm{RH}_i}\bigg|_{T=\text{const.}} &&= e_s\, K_i^e \\
K_i^{\mathrm{RH},T} &= \frac{\partial \mathrm{OLR_c}}{\partial \mathrm{RH}_i}\bigg|_{e=\text{const.}} &&= -\frac{e_s}{\mathrm{RH}}\left(\frac{de_s}{dT}\right)^{-1} K_i^T .
\end{aligned} \tag{8}
$$

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 $\mathbf{K}^{\mathrm{RH,e}}$ this factor is equal to the saturation water vapor pressure e_s. For $\mathbf{K}^{\mathrm{RH,T}}$ the dependence of e_s on T given by the Clausius Clapeyron relation has to be taken into account. $\mathbf{K}^{\mathrm{RH,e}}$ and $\mathbf{K}^{\mathrm{RH,T}}$ 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 $\mathbf{K}^{\mathrm{RH,e}}$ (Figure A.15c) are more accurate than those approximated using $\mathbf{K}^{\mathrm{RH,T}}$ (Figure A.18c). Therefore, for the further analysis we concentrate on $\mathbf{K}^{\mathrm{RH,e}}$. Overall, $\mathrm{OLR_c}$ anomalies approximated from RH anomalies agree well with true (directly calculated) $\mathrm{OLR_c}$ anomalies (Figure A.15c) and the inter-model standard deviation $\sigma(\mathrm{OLR_c})$ is well reproduced (Figure A.15d). In A.6.2 we elaborate more on the accuracy of the approximation for individual models as well as on the differences between $\mathbf{K}^{\mathrm{RH,e}}$ and $\mathbf{K}^{\mathrm{RH,T}}$.

A.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 $\mathbf{K}^{\mathrm{RH,e}}$ (Equation 6). $\mathbf{K}^{\mathrm{RH,e}}$ is negative throughout the tropical troposphere (Figure A.15a), 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 $\mathbf{K}^{\mathrm{RH,e}}$ 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 and Fueglistaler, 2020), which is why $\mathbf{K}^{\mathrm{RH,e}}$ 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 and 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

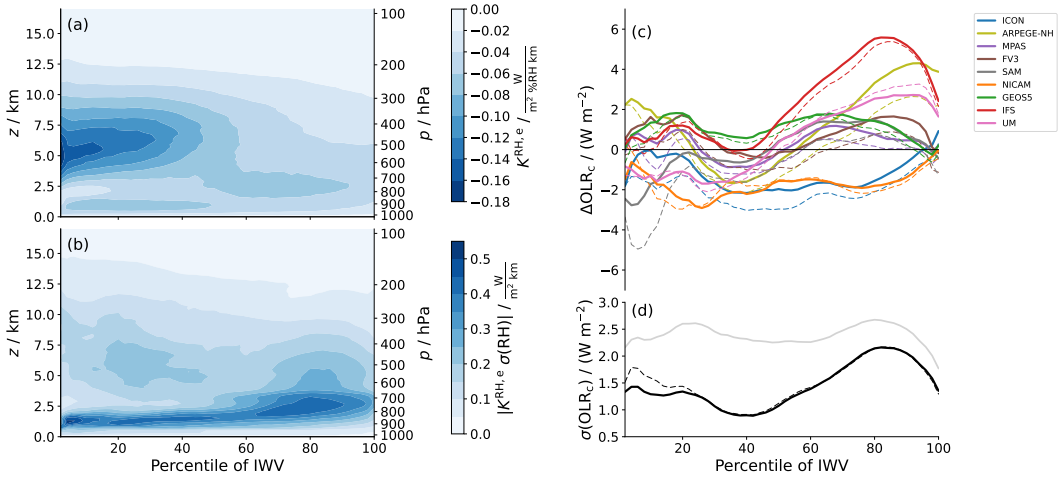


Figure A.15: Impact of RH differences on clear-sky OLR in moisture space. (a) RH response kernel $\mathbf{K}^{\mathrm{RH,e}}$ showing the sensitivity of clear-sky OLR to a 1% RH change in a 1 km layer under constant temperature for 50 blocks in moisture space, (b) inter-model standard deviation σ(RH) 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 A.14). The vertical integral of (b) is shown as the grey line in (c).

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 $\mathbf{K}^{\mathrm{RH,e}}$ 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 and 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 8). Hence, it decreases with altitude, but is approximately constant throughout moisture space. Other studies apply equal fractional perturbations in e (Shine and Sinha, 1991) or keep RH constant under a uniform temperature perturbation (Held and 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 $\mathbf{K}^{\mathrm{RH,e}}$ 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 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 $\mathbf{K}^{RH,e}$ and the RH inter-model standard deviation $\sigma(RH)$ (Figure A.15b) 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 A.15a), but the inter-model differences in RH are large (Figure A.13b) 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 A.13b) and a decreasing sensitivity of OLR_c from dry to moist regimes (Figure A.15a). 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 A.13b) 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 $\mathbf{K}^{RH,e}$ and $\sigma(RH)$, shown as the grey line in Figure A.15d, 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 A.9 and Figure A.16). This information is not contained in $\sigma(RH)$. 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 A.15d and evident from Figure A.16. 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.

A.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 (Ingram, 2010; Nakajima et al., 1992).

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 for 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 inter-comparison 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 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 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 $\sigma(\mathrm{RH})$ 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 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 $\sigma(\mathrm{RH})$ 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, which might be indicative of model differences in convective behavior. The large RH differences cause the inter-model spread in clear-sky OLR 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 $\sigma(\mathrm{RH})$ is 1–3% RH and hence less than half of the standard deviation in moist regimes. However, the sensitivity 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 local maximum in the dry regimes. This 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 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.

A.6 APPENDIX

A.6.1 *RH anomalies in individual models*

In Section A.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 A.16). 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 A.9) 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 A.4.2 and A.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 A.16d). Due to the compensation of these opposite effects the $\mathrm{OLR_c}$ anomaly in these regions is rather small (Figure A.14). In the UM model the lower and mid free troposphere are anomalously moist in dry regimes and anomalously dry in moist regimes (Figure A.16j). The resulting $\mathrm{OLR_c}$ anomalies almost fully compensate in the tropical mean (Figure A.14).

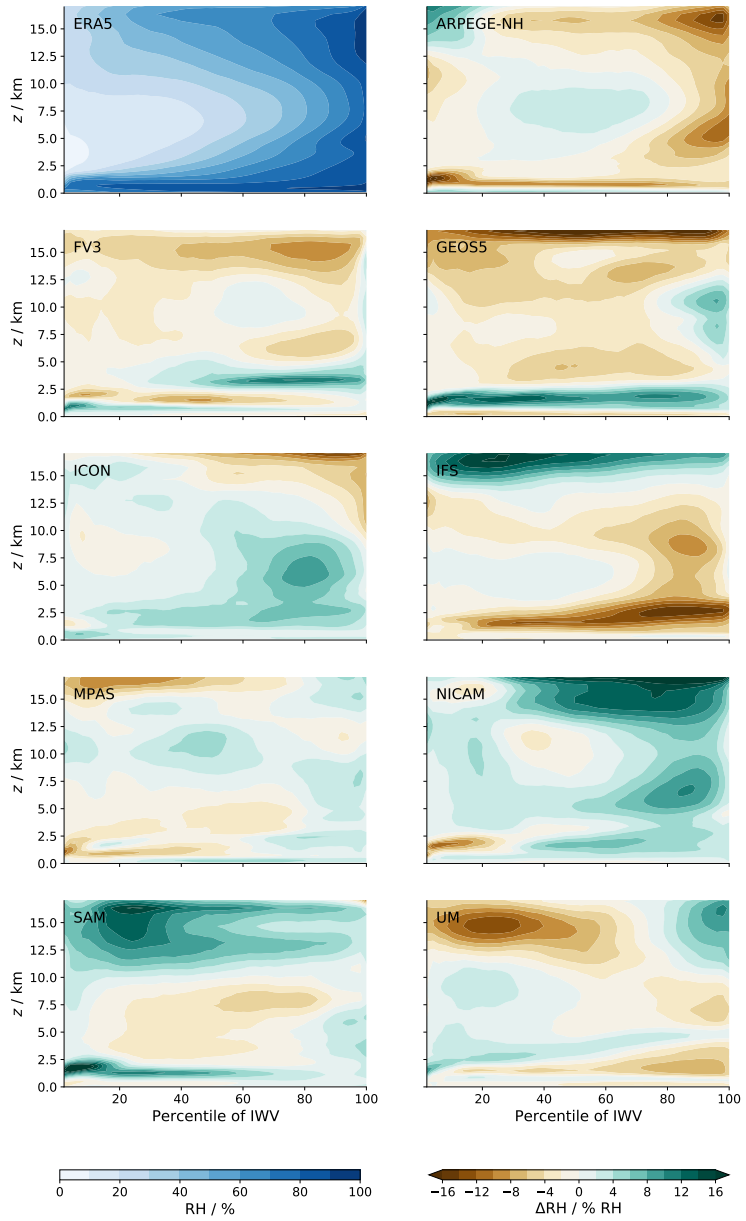


Figure A.16: RH anomalies of DYAMOND models in moisture space. The upper left panel shows the ERA5 RH distribution in moisture space, remaining panels show the deviation from the ERA5 RH for each model.

A.6.2 *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, OLR_c is calculated for the averaged ERA5 profiles in this block using the setup described in Section A.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 6) in the DYAMOND models with good accuracy (Figure A.17e). The approximation is least accurate for the NICAM model. NICAM is the model with the largest anomalies in absolute humidity (Figure A.10), 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 6. 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.

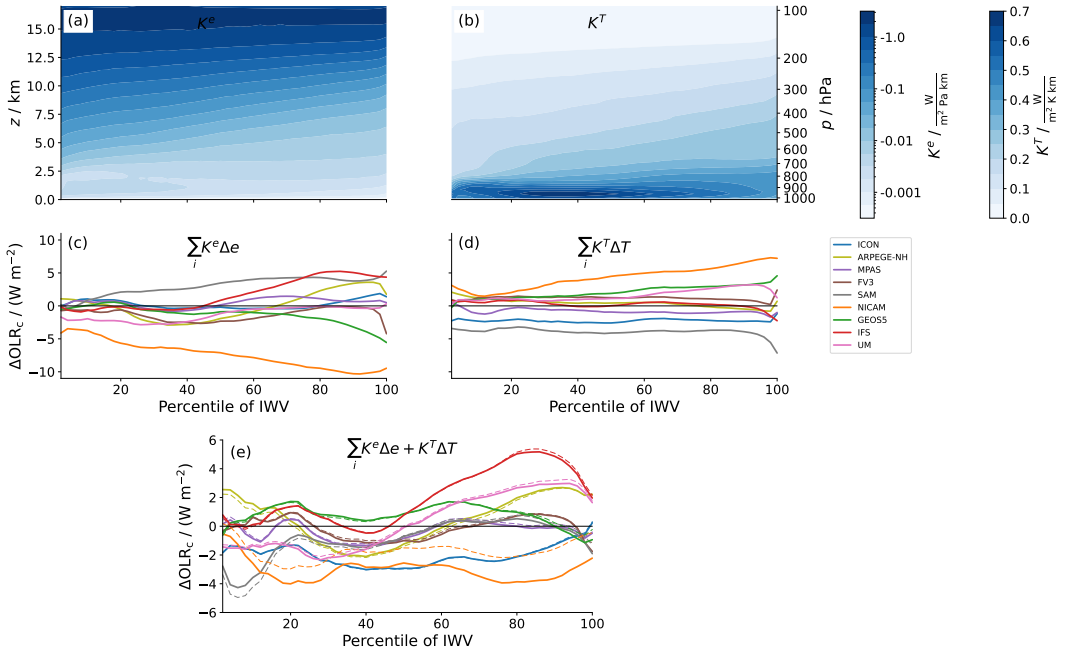


Figure A.17: Clear-sky OLR anomalies in the DYAMOND models approximated with the kernel method. (a) Water vapor response kernel $\mathbf{K^e}$ showing the sensitivity of clear-sky OLR to a change of 1 Pa in water vapor pressure e in a 1 km layer. Note the logarithmic colour scale. (b) Temperature response 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.

As explained in Section A.4.3, anomalies in $\mathrm{OLR_c}$ can also be approximated from RH anomalies and a RH kernel (Equation 6). There are two ways to define a RH kernel by varying either e or T (Equation 8), which we refer to as $\mathbf{K}^{\mathrm{RH,e}}$ and

$\mathbf{K}^{\mathrm{RH,T}}$, respectively. Our main analysis is based on $\mathbf{K}^{\mathrm{RH,e}}$ 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 A.15c). The inter-model standard deviation $\sigma(\mathrm{OLR})$ is well reproduced with the approximated $\mathrm{OLR_c}$ (Figure A.15d), 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 A.17). 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 A.18 shows $\mathbf{K}^{\mathrm{RH,T}}$ and the $\mathrm{OLR_c}$ anomalies approximated using this version of the RH kernel. $\mathbf{K}^{\mathrm{RH,T}}$ takes on larger absolute values than $\mathbf{K}^{\mathrm{RH,e}}$ (cf. Figure A.15a, note the different colour scales in Figures A.15 and A.18), 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 $\mathbf{K}^{\mathrm{RH,T}}$ is lower than in $\mathbf{K}^{\mathrm{RH,e}}$. 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 A.4.3).

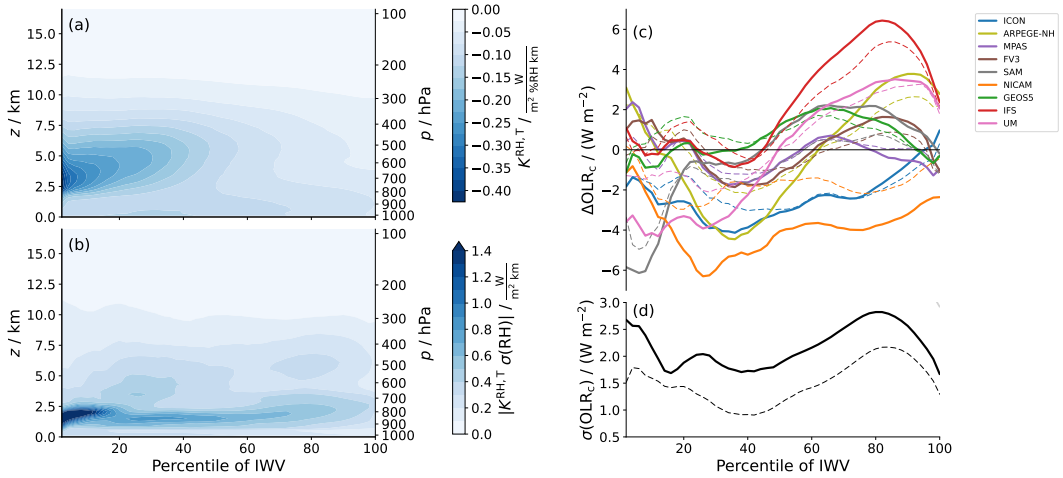


Figure A.18: As Figure A.15 but based on $\mathbf{K}^{\mathrm{RH,T}}$. Note that the colour scale in (a) and (b) is different from Figure A.15 since $\mathbf{K}^{\mathrm{RH,T}}$ takes on more negative values than $\mathbf{K}^{\mathrm{RH,e}}$.

ACKNOWLEDGEMENTS

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. et al. (1999). "The dependence of clear-sky outgoing long-wave radiation on surface temperature and relative humidity." In: *Quarterly Journal of the Royal Meteorological Society* 125.558, pp. 2103–2126. DOI: 10.1002/qj.49712555809.

Becker, Tobias and Allison A. Wing (2020). "Understanding the Extreme Spread in Climate Sensitivity within the Radiative-Convective Equilibrium Model Intercomparison Project." In: *Journal of Advances in Modeling Earth Systems* 12.10. DOI: 10.1029/2020MS002165.

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

Bony, S. et al. (2006). "How Well Do We Understand and Evaluate Climate Change Feedback Processes?" In: *Journal of Climate* 19.15, pp. 3445–3482. DOI: 10.1175/JCLI3819.1.

Bourdin, S., L. Kluft, and B. Stevens (2021). "Dependence of Climate Sensitivity on the Given Distribution of Relative Humidity." In: *Geophysical Research Letters*. DOI: https://doi.org/10.1029/2021GL092462.

Bretherton, C. S., P. N. Blossey, and M. Khairoutdinov (2005). "An Energy-Balance Analysis of Deep Convective Self-Aggregation above Uniform SST." In: *Journal of the Atmospheric Sciences* 62.12, pp. 4273–4292. DOI: 10.1175/JAS3614.1.

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

Clough, Shepard A., Michael J. Iacono, and Jean-Luc Moncet (1992). "Line-by-line calculations of atmospheric fluxes and cooling rates: Application to water vapor." In: *Journal of Geophysical Research* 97.D14, p. 15761. DOI: 10.1029/92JD01419.

Collins, M. et al. (2013). "Climate change 2013: the physical science basis." In.

Derbyshire, S. H. et al. (2004). "Sensitivity of moist convection to environmental humidity." In: *Quarterly Journal of the Royal Meteorological Society* 130.604, pp. 3055–3079. DOI: 10.1256/qj.03.130.

Dessler, A. E. and S. C. Sherwood (2000). "Simulations of tropical upper tropospheric humidity." In: *Journal of Geophysical Research* 105.D15, pp. 20155–20163. DOI: 10.1029/2000JD900231.

ECMWF (2018). "IFS Documentation CY45R1." In: chap. Part IV : Physical processes. URL: https://www.ecmwf.int/node/18714.

Hartmann, D. L. and K. Larson (2002). "An important constraint on tropical cloud - climate feedback." In: *Geophysical Research Letters* 29.20, pp. 12–1–12–4. DOI: 10.1029/2002GL015835.

Held, I. M. and B. J. Soden (2000). "Water Vapour Feedback and Global Warming." In: *Annual Review of Energy and the Environment* 25.1, pp. 441–475. DOI: 10.1146/annurev.energy.25.1.441.

Hersbach, Hans et al. (2020). "The ERA5 global reanalysis." In: *Quarterly Journal of the Royal Meteorological Society* 146.730, pp. 1999–2049. DOI: 10.1002/qj.3803.

Ingram, William (2010). "A very simple model for the water vapour feedback on climate change." In: *Quarterly Journal of the Royal Meteorological Society* 136.646, pp. 30–40. DOI: 10.1002/qj.546.

Jeevanjee, Nadir and Stephan Fueglistaler (2020). "Simple Spectral Models for Atmospheric Radiative Cooling." In: *Journal of the Atmospheric Sciences* 77.2, pp. 479–497. DOI: 10.1175/JAS-D-18-0347.1.

Jensen, Eric J. et al. (1999). "High humidities and subvisible cirrus near the tropical tropopause." In: *Geophysical Research Letters* 26.15, pp. 2347–2350. DOI: 10.1029/1999GL900266.

John, V. O. and B. J. Soden (2007). "Temperature and humidity biases in global climate models and their impact on climate feedbacks." In: *Geophysical Research Letters* 34.18, p. L18704. DOI: 10.1029/2007GL030429.

John, V. O. et al. (2006). "Understanding the variability of clear-sky outgoing longwave radiation based on ship-based temperature and water vapour measurements." In: *Quarterly Journal of the Royal Meteorological Society* 132.621, pp. 2675–2691. DOI: 10.1256/qj.05.70.

Koll, Daniel D. B. and Timothy W. Cronin (2018). "Earth's outgoing longwave radiation linear due to H2O greenhouse effect." In: *Proceedings of the National Academy of Sciences* 115.41, pp. 10293–10298. DOI: 10.1073/pnas.1809868115.

Luo, Z. and W. B. Rossow (2004). "Characterizing Tropical Cirrus Life Cycle, Evolution, and Interaction with Upper-Tropospheric Water Vapor Using Lagrangian Trajectory Analysisof Satellite Observations." In: *Journal of Climate* 17.23, pp. 4541–4563. DOI: 10.1175/3222.1.

Mapes, Brian E. et al. (2018). "The Meandering Margin of the Meteorological Moist Tropics." In: *Geophysical Research Letters* 45.2, pp. 1177–1184. DOI: 10.1002/2017GL076440.

Miura, H. et al. (2007). "A Madden-Julian Oscillation Event Realistically Simulated by a Global Cloud-Resolving Model." In: *Science* 318.5857, pp. 1763–1765. DOI: 10.1126/science.1148443.

Mlawer, Eli J. et al. (1997). "Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave." In: *Journal of Geophysical Research: Atmospheres* 102.D14, pp. 16663–16682. DOI: 10.1029/97JD00237.

Monteiro, Joy Merwin, Jeremy McGibbon, and Rodrigo Caballero (2018). "sympl (v. 0.4.0) and climt (v. 0.15.3) – towards a flexible framework for building model hierarchies in Python." In: *Geoscientific Model Development* 11.9, pp. 3781–3794. DOI: 10.5194/gmd-11-3781-2018.

Nakajima, Shinichi, Yoshi-Yuki Hayashi, and Yutaka Abe (1992). "A Study on the "Runaway Greenhouse Effect" with a One-Dimensional Radiative–Convective Equilibrium Model." In: *Journal of the Atmospheric Sciences* 49.23, pp. 2256–2266. DOI: 10.1175/1520-0469(1992)049<2256:ASOTGE>2.0.CO;2.

Naumann, A. K. and C. Kiemle (2020). "The vertical structure and spatial variability of lower-tropospheric water vapor and clouds in the trades." In: *Atmospheric Chemistry and Physics* 20.10, pp. 6129–6145. DOI: 10.5194/acp-20-6129-2020.

Petty, G. W. (2006). *A First Course in Atmospheric Radiation*. Sundog Pub. ISBN: 0972903313.

Pierrehumbert, R. T., H. Brogniez, and R. Roca (2007). "The Global Circulation of the Atmosphere." In: ed. by T. Schneider and A. Sobel. Princeton University Press. Chap. On the relative humidity of the atmosphere, pp. 143–185.

Pierrehumbert, R. T. and R. Roca (1998). "Evidence for Control of Atlantic Subtropical Humidity by Large Scale Advection." In: *Geophysical Research Letters* 25.24, pp. 4537–4540. DOI: 10.1029/1998GL900203.

Pincus, Robert et al. (2015). "Radiative flux and forcing parameterization error in aerosol-free clear skies." In: *Geophysical Research Letters* 42.13, pp. 5485–5492. DOI: 10.1002/2015GL064291.

Po-Chedley, Stephen et al. (2018). "Sources of Intermodel Spread in the Lapse Rate and Water Vapor Feedbacks." In: *Journal of Climate* 31.8, pp. 3187–3206. DOI: 10.1175/JCLI-D-17-0674.1.

Po-Chedley, S. et al. (2019). "Climatology Explains Intermodel Spread in Tropical Upper Tropospheric Cloud and Relative Humidity Response to Greenhouse Warming." In: *Geophysical Research Letters* 46.22, pp. 13399–13409. DOI: 10.1029/2019GL084786.

Retsch, M. H., T. Mauritsen, and C. Hohenegger (2019). "Climate Change Feedbacks in Aquaplanet Experiments With Explicit and Parametrized Convection for Horizontal Resolutions of 2,525 Up to 5 km." In: *Journal of Advances in Modeling Earth Systems* 11.7, pp. 2070–2088. DOI: 10.1029/2019MS001677.

Romps, D. M. (2014). "An Analytical Model for Tropical Relative Humidity." In: *Journal of Climate* 27.19, pp. 7432–7449. DOI: 10.1175/JCLI-D-14-00255.1.

Satoh, M. et al. (2019). "Global Cloud-Resolving Models." In: *Current Climate Change Reports* 5.3, pp. 172–184. DOI: 10.1007/s40641-019-00131-0.

Schulz, H. and B. Stevens (2018). "Observing the Tropical Atmosphere in Moisture Space." In: *Journal of the Atmospheric Sciences* 75.10, pp. 3313–3330. DOI: 10.1175/JAS-D-17-0375.1.

Schulzweida, Uwe (2019). *CDO User Guide*. en. DOI: 10.5281/zenodo.3539275.

Sherwood, S. C. (1996). "Maintainance of the Free-Tropospheric Tropical Water Vapor Distribution. Part II: Simulation by Large-Scale Advection." In: *Journal of Climate* 9.11, pp. 2919–2934. DOI: 10.1175/1520-0442(1996)009<2919:MOTFTT>2.0.CO;2.

Sherwood, S. C. et al. (2010). "Relative humidity changes in a warmer climate." In: *Journal of Geophysical Research* 115.D9, p. D09104. DOI: 10.1029/2009JD012585.

Shine, Keith P. and Ashok Sinha (1991). "Sensitivity of the Earth's climate to height-dependent changes in the water vapour mixing ratio." In: *Nature* 354.6352, pp. 382–384. DOI: 10.1038/354382a0.

Simpson, G. C. (1928). "Some studies in terrestrial radiation." In: *Memoirs of the Royal Meteorological Society* 2.16, pp. 69–95.

Soden, B. J. and I. M. Held (2006). "An Assessment of Climate Feedbacks in Coupled Ocean–Atmosphere Models." In: *Journal of Climate* 19.14, pp. 3354–3360. DOI: 10.1175/JCLI3799.1.

Soden, B. J. et al. (2008). "Quantifying Climate Feedbacks Using Radiative Kernels." In: *Journal of Climate* 21.14, pp. 3504–3520. DOI: 10.1175/2007JCLI2110.1.

Spencer, R. W. and W. D. Braswell (1997). "How Dry is the Tropical Free Troposphere? Implications for Global Warming Theory." In: *Bulletin of the Ameri-*

can Meteorological Society 78.6, pp. 1097–1106. DOI: 10.1175/1520-0477(1997)078<1097:hdittf>2.0.co;2.

Stevens, B. et al. (2017). "Structure and Dynamical Influence of Water Vapor in the Lower Tropical Troposphere." In: *Surveys in Geophysics* 38.6, pp. 1371–1397. DOI: 10.1007/s10712-017-9420-8.

Stevens, Bjorn et al. (2019). "DYAMOND: the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains." In: *Progress in Earth and Planetary Science* 6.61. DOI: 10.1186/s40645-019-0304-z.

Stevens, Bjorn et al. (2020). "The Added Value of Large-eddy and Storm-resolving Models for Simulating Clouds and Precipitation." In: *Journal of the Meteorological Society of Japan. Ser. II* 98.2, pp. 395–435. DOI: 10.2151/jmsj.2020-021.

Taylor, Karl E., Ronald J. Stouffer, and Gerald A. Meehl (2012). "An Overview of CMIP5 and the Experiment Design." In: *Bulletin of the American Meteorological Society* 93.4, pp. 485–498. DOI: 10.1175/BAMS-D-11-00094.1.

Tomita, H. et al. (2005). "A global cloud-resolving simulation: Preliminary results from an aqua planet experiment." In: *Geophysical Research Letters* 32.8. DOI: 10.1029/2005GL022459.

Vial, J., J.-L. Dufresne, and S. Bony (2013). "On the interpretation of inter-model spread in CMIP5 climate sensitivity estimates." In: *Climate Dynamics* 41.11-12, pp. 3339–3362. DOI: 10.1007/s00382-013-1725-9.

Vömel, H. (2002). "Balloon-borne observations of water vapor and ozone in the tropical upper troposphere and lower stratosphere." In: *Journal of Geophysical Research* 107.D14. DOI: 10.1029/2001JD000707.

Wing, A. A. et al. (2020). "Clouds and Convective Self-Aggregation in a Multimodel Ensemble of Radiative-Convective Equilibrium Simulations." In: *Journal of Advances in Modeling Earth Systems* 12.9. DOI: 10.1029/2020MS002138.

Wing, Allison A. et al. (2017). "Radiative-Convective Equilibrium Model Intercomparison Project." In: *Geoscientific Model Development*. DOI: 10.5194/gmd-2017-213.

Xue, Y. et al. (2020). "Assessment of Upper Tropospheric Water Vapor Monthly Variation in Reanalyses With Near-Global Homogenized 6.5-μm Radiances From Geostationary Satellites." In: *Journal of Geophysical Research: Atmospheres* 125.18. DOI: 10.1029/2020JD032695.

B

HOW MODEL UNCERTAINTIES INFLUENCE TROPICAL HUMIDITY IN GLOBAL STORM-RESOLVING SIMULATIONS

The work in this is under review in the *Journal of Advances in Modeling Earth Systems* and was submitted as:

Lang, T., Naumann, A. K., Buehler, S. A., Stevens, B., Schmidt, H. and Aemisegger, F. (2022). How model uncertainties influence tropical humidity in global storm-resolving simulations.

A preprint is available at `https://www.essoar.org/doi/abs/10.1002/essoar.10512585.1`

CONTRIBUTIONS I conceptualized the study. A. K. Naumann, S. Buehler and B. Stevens supported the conceptualization. A. K. Naumann and S. A. Buehler supervised the study. H. Schmidt performed three of the model simulations. I performed the remaining five simulations and trajectory calculations, conducted the analysis and prepared the draft manuscript. All co-authors contributed to the discussion of the results and the manuscript at all stages.

How model uncertainties influence tropical humidity in global storm-resolving simulations

Theresa Lang[1,2], Ann Kristin Naumann[3,1], Stefan A. Buehler[1], Bjorn Stevens[3], Hauke Schmidt[3], Franziska Aemisegger[4]

[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*

[4]*Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland*

KEY POINTS

- Sensitivity experiments suggest that parameterizations are the major source of relative humidity spread across global storm-resolving models
- Vertical mixing processes strongly impact the humidity of the moist tropics by affecting last-saturation statistics within the tropics
- The humidity of the dry tropics is disproportionately sensitive to changes in the pathways of exchange with the extra-tropics

ABSTRACT

We conduct a series of eight 45-day experiments with a global storm-resolving model (GSRM) to test the sensitivity of relative humidity $\mathcal{R}$ in the tropics to changes in model resolution and parameterizations. These changes include changes in horizontal and vertical grid spacing as well as in the parameterizations of microphysics and turbulence, and are chosen to capture currently existing differences among GSRMs. To link the $\mathcal{R}$ distribution in the tropical free troposphere with processes in the deep convective regions, we adopt a trajectory-based assessment of the last-saturation paradigm. The perturbations we apply to the model result in tropical mean $\mathcal{R}$ changes ranging from 0.5% to 8% (absolute) in the mid troposphere. The generated $\mathcal{R}$ spread is similar to that in a multi-model ensemble of GSRMs and smaller than the spread across conventional general circulation models, supporting that an explicit representation of deep convection reduces the

uncertainty in tropical $\mathcal{R}$. The largest $\mathcal{R}$ changes result from changes in parameterizations, suggesting that model physics represent a major source of humidity spread across GSRMs. The $\mathcal{R}$ in the moist tropical regions is disproportionately sensitive to vertical mixing processes within the tropics, which impact $\mathcal{R}$ through their effect on the last-saturation temperature rather than their effect on the evolution of the humidity since last-saturation. In our analysis the $\mathcal{R}$ of the dry tropical regions strongly depends on the exchange with the extra-tropics. The interaction between tropics and extratropics could change with warming and presage changes in the radiatively sensitive dry regions.

PLAIN LANGUAGE SUMMARY

Water vapor is the most important greenhouse gas in the atmosphere. Therefore, for the prediction of future warming it is important that climate models capture the distribution of atmospheric humidity and its change under warming. However, climate models currently strongly disagree in their representation of humidity, causing uncertainty in climate predictions. A recent study has shown that, while there is better agreement among the newest generation of climate models, so called global storm-resolving models, the remaining inter-model differences are still relevant and therefore need to be better understood. To narrow down the causes of these differences, in this study we examine how much the humidity in a storm-resolving model changes in response to changes in different model components, which are chosen to reflect the differences that currently exist between models. We find the largest humidity changes in response to changes in the model's representation of sub-grid scale processes. In storm-resolving models these are turbulent motions and cloud microphysics. Our results suggest that differences in the representation of these processes cause a major part of the humidity differences between storm-resolving models.

B.1 INTRODUCTION

The aim of this study is to better understand sources of uncertainties in modelling processes that drive the distribution of tropical free-tropospheric relative humidity. Therefore, we examine how much and through which physical mechanisms the relative humidity in a global storm-resolving model (GSRM) – the newest generation of climate models with high horizontal resolution and explicit simulation of convection – is affected by changes in model resolution and paramtererizations.

Free-tropospheric relative humidity plays an important role in determining Earth's climate sensitivity. The combined effect of the water vapor and lapse rate feedbacks – the two most important feedbacks acting under clear-sky conditions – largely depends on how relative humidity responds to warming (Held and Shell, 2012). While to first order relative humidity is expected to stay constant under warming (Held and Soden, 2000), even small deviations from this constancy significantly impact the clear-sky feedback by altering the cancellation between water vapor and lapse rate feedbacks in the saturated parts of the emission spectrum

(Bony et al., 2006). In line with that, model differences in the relative humidity response control the prevailing spread in clear-sky feedback across general circulation models (GCMs; Vial et al., 2013). Since the relative humidity change simulated by GCMs is described by an upward shift following the rising isotherms, differences in the models'relative humidity response are closely related to differences in their climatology (Po-Chedley et al., 2019). Even if relative humidity does not change with warming, its present-day value might affect the clear-sky feedback. While no systematic relationship between present-day state and feedbacks has been found for GCMs (John and Soden, 2007), 1D radiative convective equilibrium (RCE) studies suggest that particularly at high surface temperatures characteristic of the tropics, the closing of the spectral window results in a strong dependence of the clear-sky feedback on relative humidity (Bourdin et al., 2021; Kluft et al., 2019; McKim et al., 2021). Thus, to develop a more fundamental understanding of climate and climate change, we will need to understand what sets the distribution of relative humidity, how it might change, and why it differs across models.

The sources of the relative humidity spread across models are poorly understood. One reason for this is the number of processes that affect humidity, many of which are poorly constrained in GCMs. In particular deep convection, the process accounting for most of the vertical transport of water vapour in the tropical atmosphere, is not resolved in these models and needs to be parameterized. An important step has been made with the development of global storm-resolving models (GSRMs; Satoh et al., 2019). With grid spacings of a few kilometers, these models simulate deep convection explicitly and thereby forego the need for convective parameterizations. At present, due to the high computational effort, storm-resolving simulations are limited to time scales of days to months. A first intercomparison of GSRMs, the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains (DYAMOND; Stevens et al., 2019) project, indicates that the inter-model spread in tropical free-tropospheric humidity is indeed reduced compared to GCMs (Lang et al., 2021). While this is a promising result and highlights the benefit of even approximately resolving deep convection, the study also showed that the remaining differences in relative humidity are still an important source of uncertainty for the clear-sky outgoing longwave radiation (OLR).

In this study, we attempt to understand the reasons behind the remaining relative humidity differences. To this end, we examine how the tropical humidity simulated by a GSRM changes in response to modifications in model resolution and model physics. These modifications are chosen to resemble currently existing differences across GSRMs. A large ensemble of back-trajectories started from the tropical mid troposphere allows us to examine the history of the air parcels arriving in these regions and hence the physical mechanisms behind humidity changes in the experiments.

To examine these mechanisms we make use of the last-saturation or advection-condensation paradigm (Sherwood, 1996; Sherwood et al., 2010), which represents

the simplest model of what determines the distribution of free-tropospheric humidity. Assuming that water vapor behaves as a conservative tracer for which condensation is a permanent sink term, the water vapor content of an air parcel is determined by its temperature at the instant at which condensation last occurred. Inside a cloud, an air parcel's specific humidity is at saturation. As the parcel rises, it looses water vapor by condensation. Outside the cloud, the air parcel subsides and warms adiabatically, while maintaining the specific humidity it had when it was last saturated, so its relative humidity decreases. The regions where last-condensation events typically occur are often referred to as the “source regions" or “origins" of free-tropospheric air. The source regions of tropical free-tropospheric air are mainly located in the tropical deep convective regions, but a significant part of the air in the dry subtropical subsidence regions also originates from the extra-tropics (e.g. Aemisegger et al., 2021; Cau et al., 2007; Roca et al., 2012). According to the last-saturation model, the relative humidity in a given target region only depends on the properties – mainly the temperature – of the source region and the target region.

Numerical implementations of the last-saturation model, which used large-scale wind and temperature fields from meteorological analyses to calculate Lagrangian back-trajectories, have been successful in reproducing the observed free-tropospheric relative humidity distribution (e.g. Dessler and Sherwood, 2000; Pierrehumbert and Roca, 1998; Sherwood, 1996). This has lead to the conclusion that the relative humidity distribution is determined by circulation and temperature structure, while any moisture sources or sinks changing the specific humidity of an air parcel after the last-saturation event are of minor importance. These sources and sinks include evaporation of cloud condensate or from precipitation, as well as mixing due to motions on scales not resolved in the wind field used for the trajectory calculation. This is not to say that these processes are unimportant, rather to say that to the extent they are important, it is through their indirect influence on the atmospheric circulation and the temperature structure, which ultimately determine the location of last-saturation events.

While the moisture sources and sinks after last-saturation appear to play a secondary role in determining spatial variations of relative humidity in the real atmosphere or a given model, it is less clear whether they might be important when it comes to explaining the more subtle humidity differences between models, particularly when different parameterizations for the processes causing the sources and sinks, i.e. microphysics and turbulence, are used. To test this, we calculate back-trajectories to perform two types of Lagrangian relative humidity reconstructions for our model experiments. The first one is an implementation of the last-saturation model and therefore only takes into account the properties of air parcels in the source and target regions. The second one additionally accounts for parameterized moisture sources and sinks during the advection of air parcels to the target region. Comparing the two types of reconstructions allows us to quantify the importance of changes in moisture sources and sinks in causing the relative humidity changes in our sensitivity experiments. To

our knowledge, the last-saturation model has neither been used to understand differences between models, nor has it been implemented based on wind fields of simulations at storm-resolving resolution. This study therefore also represents a test of how useful the last-saturation model is in explaining differences between models as they begin to resolve the spectrum of vertical motions in the atmosphere.

This paper is organized as follows: Section B.2 describes the model setup and the sensitivity experiments performed. In Section B.3 the humidity changes produced in our sensitivity experiments are shown and discussed. The Lagrangian relative humidity reconstructions based on back-trajectories are introduced in Section B.4. Section B.5 presents insights on the mechanisms behind the humidity changes from the last-saturation model.

B.2 MODEL AND EXPERIMENTS

To examine how changes in model parameterizations and model resolution affect tropical relative humidity in a GSRM, we run a series of sensitivity experiments with the ICOsahedral Nonhydrostatic model (ICON; Zängl et al., 2015) in its storm-resolving "Sapphire" configuration (Hohenegger et al., 2022) with prescribed sea surface temperature (SST).

B.2.1 *Control experiment*

The control experiment is run with a quasi-uniform horizontal grid spacing of 5 km. For the analysis, the model output is interpolated from the native icosahedral ICON grid to a regular $0.1^\circ \times 0.1^\circ$ latitude-longitude grid. The vertical grid consists of 110 hybrid sigma height levels between the surface and a height of 75 km. Over a flat surface at sea level, the distance between model levels in the free troposphere (between about 8 km to 19 km) is constant at 400m, gradually decreasing towards the surface and increasing towards the model top. The model time step is 40 seconds. For the treatment of microphysical processes, a one-moment scheme with five hydrometeor categories as described by Baldauf et al., 2011 is used. Turbulent mixing is represented by a classical 3D Smagorinksy scheme (Smagorinsky, 1963) with the modification by Lilly, 1962 to account for thermal stratification (Dipankar et al., 2015). Radiative transfer is calculated at every grid point every 15 minutes using the RTE-RRTMGP scheme (Pincus et al., 2019). The JSBACH land model (Raddatz et al., 2007) is used to represent the physical properties of the land surface and land-atmosphere interactions.

The experimental protocol of our experiments closely follows that specified by the DYAMOND inter-model comparison (Stevens et al., 2019), with initial conditions taken from the global (9km) analysis by the European Centre of Medium Range Weather Forecast (ECMWF). After initialization, the simulations run freely without further forcing. ECMWF operational daily SST and sea-ice concentration are used as boundary conditions. The simulations start at 0 UTC on

27 June 2021 and span a time period of 45 days.

To test the extent to which humidity differences in our 45-day simulations might reflect sampling error, we perform a second control experiment (Control 2) with perturbed initial conditions. While the boundary conditions are kept identical to those in the control run, the atmosphere is initialised from the ECMWF analysis for 0 UTC on 28 June 2021, i.e. one day later than in the control experiment.

B.2.2 *Sensitivity experiments*

The changes we apply in our sensitivity experiments are chosen to resemble differences in model configuration across the DYAMOND models (Stevens et al., 2019), which reflect current differences in modeling approaches between modeling groups. The DYAMOND models differ in various aspects of their configuration. On the one hand, they differ in the design of their dynamical core. While (with the exception of two models) they agree on the equations they solve (fully-compressible Navier-Stokes equations), they differ in their numerical grids and the numerical methods they use to solve the equations. This not only influences their "effective" resolution, but also conditions the behavior of the parameterizations which act on the grid scale. On the other hand, the models differ in the parameterizations they use to represent the effects of subgrid-scale processes. For the sensitivity experiments we have to concentrate on a subset of these differences that can be tested with the ICON model. We attempt to cover the different types of uncertainties by examining the sensitivity of relative humidity to the model resolution as well as two different parameterizations. Our sensitivity experiments are described in the following and summarized in Table 1.

Even if at 5 km most of the energy in the spectrum of vertical motions is resolved, the updrafts of most deep convective systems remain poorly resolved or aliased to larger scales. To test the extent to which relative humidity is affected by changes in model resolution we perform three experiments. In the $\Delta x/2$ experiment the horizontal grid spacing is halved relative to the control experiment to 2.5 km. For the $2\Delta z$ and $\Delta z/2$ experiments the number of vertical levels is decreased to 55 and increased to 190, respectively. This results in a doubling and halving of the vertical grid spacing in the free troposphere relative to the control experiment to 800 m and 200 m, respectively. Note that by GSRM standards (if not by GCM standards) a vertical grid spacing of 800 m is exceptionally coarse and was not employed in any of the DYAMOND models.

In three further experiments we test the sensitivity of relative humidity to changes in the parameterizations of turbulence and microphysics. These parameterizations contain a large number of tunable parameters and we do not attempt to systematically test the sensitivity to all of them. Instead we focus on contrasting models, which we see as a more extreme case than parameter

sensitivities, although in one case we also explore a common parameter sensitivity.

Storm-resolving models typically use turbulence paramterizations that are not well adapted to global simulations at kilometer-scales. On the one hand, regional storm-resolving models have often used turbulence closures designed for LES simulations (like the Smagorinsky-Lilly scheme used in our control simulation), although the underlying assumption that the truncation scale lies within the inertial range of three-dimensional homogeneous and isotropic turbulence (Lilly, 1967) is not satisfied at storm-resolving scales (e.g. Bryan et al., 2003). On the other hand, many of the global DYAMOND models employed turbulence schemes that were inherited from their coarser-resolution predecessors. Similarly, the storm-resolving version of the ICON model was run with a total turbulent energy (TTE) scheme (Mauritsen et al., 2007) that was originally used at much coarser resolutions in the early stages of its development (Mauritsen et al., 2022). To examine the impact of different turbulence parameterizations on relative humidity, we exchange the Smagorinsky scheme used in the control simulation with this TTE scheme. The two schemes differ in several aspects. The Smagorinsky scheme calculates both vertical and horizontal mixing of momentum and scalar variables (although we find that horizontal mixing tendencies of specific humidity are negligible compared to vertical mixing tendencies at 5 km horizontal resolution, see also Section B.4.4). The exchange coefficients are specified using a mixing length scale that depends on height and the model grid spacing, the 3D wind shear and static stability. The TTE scheme, on the other hand, only represents vertical mixing. The turbulent exchange coefficients are specified using a height-dependent mixing length scale and a velocity scale. The latter is determined from a prognostic equation for TTE that takes into account shear production, dissipation, third-order flux divergence and buoyancy production, which allows for mixing in more stably stratified situations than in the ICON implementation of the Smagorinsky-Lilly model.

To test the sensitivity of relative humidity to the microphysics parameterization, in the 2-mom experiment we exchange the one-moment scheme with the two-moment scheme by Seifert and Beheng, 2001. While the DYAMOND models all use one-moment schemes, this mainly reflects the consensus that the scheme should be computationally efficient. The degree of complexity required in the cloud microphysics is an open question, and more complex two-moment schemes have also been proposed for storm-resolving simulations (e.g. Morrison et al., 2005; Phillips et al., 2007). The one-moment and two-moment microphysics implemented in ICON differ in many of their parameters, so changes emerging in the 2-mom experiment do not only result from the fact that two moments instead of one moment of the particle size distributions are predicted.

In an additional microphysics experiment, the $2v_{\text{ice}}$ experiment, we perturb the one-moment microphysics by increasing the terminal fall speed of ice particles v_{ice},

Table 1: Summary of simulations performed with the ICON model.

Name	Description
Control	Control simulation with 5 km horizontal grid spacing, 110 vertical levels (400 m grid spacing in the free troposphere), three-dimensional Smagorinsky turbulence and one-moment microphysics
Control 2	As Control, but with perturbed initial conditions to estimate internal variability
$\Delta x/2$	Horizontal grid spacing halved to 2.5 km
$2\Delta z$	Number of vertical levels reduced to 55 (800 m grid spacing in the free troposphere)
$\Delta z/2$	Number of vertical levels increased to 190 (200 m grid spacing in the free troposphere)
TTE	Turbulence scheme exchanged by a one-dimensional total turbulent energy (TTE) scheme
2-mom	Microphysics scheme exchanged by a two-moment scheme
$2v_{\text{ice}}$	Increased (approximately doubled) fall speed of ice particles in the one-moment microphysics

which represents a common tuning parameter. In the one-moment scheme it is parameterized as a function of ice mass mixing ratio q_{ice} and air density ρ:

$$v_{\text{ice}} = a(\rho q_{\text{ice}})^b (\rho_0/\rho)^c \tag{9}$$

with $\rho_0 = 1.225\,\text{kg}\,\text{m}^{-2}$ is the air density at surface conditions. The parameters a, b and c are set to 1.25, 0.16 and 0.33, respectively. For our sensitivity experiment we increase a to 3.29, which corresponds to the value originally proposed by Heymsfield and Donner, 1990, and c to 0.4, thereby moving it closer to the value of 0.5 used in the two-moment scheme of ICON. Combined, these changes approximately double the fall speed of ice particles for a given q_i and ρ.

B.3 SENSITIVITY OF RELATIVE HUMIDITY TO CHANGES IN MODEL RESOLUTION AND PARAMETERIZATIONS

Figure B.19 shows how the tropical mean vertical profile of relative humidity changes in our sensitivity experiments. Here, relative humidity $\mathcal{R}$ is calculated as

$$\mathcal{R} = \frac{q}{q^*(T,p)} \tag{10}$$

with the specific humidity q and the saturation specific humidity $q^* = \frac{\frac{M_w}{M_d} e^*(T)}{p-(1-\frac{M_w}{M_d})e^*(T)}$, where e^* is the saturation water vapor pressure at temperature T, p is the pressure and M_w and M_d are the molar masses of water vapor

and dry air, respectively. For e^* we take the value with respect to water for T above the triple point of water T_t and the value with respect to ice for T below $T_t - 23\,\mathrm{K}$. For intermediate T a combination of both is used following the documentation of the Integrated Forecast System (ECMWF, 2018). Note that a more common definition of relative humidity uses saturation water vapor pressure instead of specific humidity. We use equation 10 to make the definition of $\mathcal{R}$ consistent with the one we use for the Lagrangian reconstructions in Section B.4. This definition is typically used in last-saturation studies (e.g. Sherwood et al., 2010) since specific humidity is the conserved quantity after last-saturation. Numerically, the difference between the two definitions is typically within 1%.

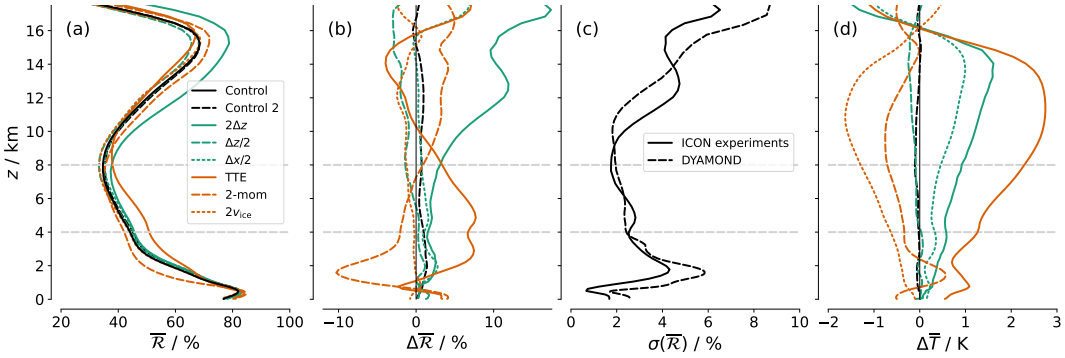


Figure B.19: Changes in tropical mean relative humidity ($\overline{\mathcal{R}}$) and temperature ($\overline{T}$) resulting from changes in model resolution and parameterizations in the sensitivity experiments. (a) Vertical profiles of $\overline{\mathcal{R}}$ in control and sensitivity experiments, (b) change in $\overline{\mathcal{R}}$ compared to the control experiment and (c) standard deviation of $\overline{\mathcal{R}}$ across ICON experiments (solid) and the DYAMOND multi-model ensemble (dashed). (d) Change in temperature $\overline{T}$ compared to the control experiment. Horizontal dashed lines mark the altitude region between 4 km and 8 km, for which the mechanisms behind the $\mathcal{R}$ changes are investigated based on back-trajectories.

First it is worth noting that the $\mathcal{R}$ spread produced by our experiments is similar to the inter-model spread in the DYAMOND ensemble (Figure B.19c). Based on the DYAMOND ensemble, Lang et al., 2021 showed that the $\mathcal{R}$ spread across GSRMs is reduced compared to classical GCMs. This is possibly related to the omission of convective parameterisations, which represent a major source of uncertainty in GCMs. Our experiments support this by showing that even strong perturbations in GSRMs do not reproduce the spread across models with convective parameterizations.

Of the experiments with changed model resolution the largest changes in $\mathcal{R}$ are seen in the $2\Delta z$ experiment with reduced vertical resolution (Figure B.19a,b). $\mathcal{R}$ increases particularly in the upper troposphere, where the difference to the control experiment exceeds 10%. In line with this, increasing the vertical resolution ($\Delta z/2$) reduces $\mathcal{R}$ in the upper troposphere. However, the magnitude of the drying is much smaller than the moistening in the $2\Delta z$ experiment, so the $\mathcal{R}$ profile shows signs of convergence at vertical resolutions around the one used in the control

experiment. Increasing the horizontal resolution ($\Delta x/2$) also only leads to a minor increase of $\mathcal{R}$ in the lower and mid troposphere. Given that the $2\Delta z$ experiment represents a rather extreme case, in the sense that GSRMs are not commonly run at such coarse vertical resolution, these results suggest that changing model resolution within the general scale of GSRM resolution does not represent a major uncertainty for $\mathcal{R}$, unless it is chosen exceptionally coarse. Note that this does not exclude the possibility that increasing resolution to even finer scales (on the order of 200 m) could make a significant difference, which cannot be tested with the chosen setup and available computational resources.

Comparably large $\mathcal{R}$ changes occur in the TTE and 2-mom experiments, in which the parameterizations of turbulence and microphysics were changed. The largest changes occur in the lower and mid troposphere, where they have a larger impact on the clear-sky OLR than those in the upper troposphere (Lang et al., 2021). Changing to the TTE turbulence scheme results in a strong increase in $\mathcal{R}$ of up to 8% over a broad altitude layer between 2 km to 6 km. This change will be examined in more detail in the following sections as part of our last-saturation analysis of the mid troposphere. Changing to the 2-mom microphysics scheme leads to a strong (up to 10%) decrease in $\mathcal{R}$ that is concentrated in a rather shallow layer between 1 km and 3 km in the lower free troposphere. Dividing the tropics into different moisture regimes also shows that this drying is concentrated in the dry subsidence regimes of the tropics, where shallow clouds prevail (not shown). This might indicate that the details in the formulation of the microphysics matter particularly in the shallow cloud regime, where humidity is not as strongly constrained by the dynamics as in deep convective regimes. Increasing the fall speed of ice particles in the 1-mom scheme ($2v_{\text{ice}}$) has a smaller effect on $\mathcal{R}$ than changing to the two-moment scheme. $\mathcal{R}$ slightly decreases in the mid to upper troposphere, whereas lower-tropospheric $\mathcal{R}$ is hardly affected. This may be expected, since ice particles mainly exist at higher altitudes with temperatures below the melting point (located at a height of about 5 km in our experiments). Changing between one- and two-moment microphysics, on the other hand, potentially affects the characteristics of all types of hydrometeors.

$\mathcal{R}$ changes in most sensitivity experiments are larger than the difference between the two control experiments (Control and Control 2) which serves as an estimate of internal variability. Exceptions are the very subtle changes in the $2v_{\text{ice}}$ and $\Delta z/2$ experiments in the lower free troposphere and in the $\Delta x/2$ experiment in the upper troposphere. We conclude that the differences we find in tropical mean $\mathcal{R}$ mostly represent systematic differences resulting from the applied perturbations rather than internal variability.

Temperature profiles differ substantially between the experiments (Figure B.19d). Temperature differences that exist in the lower troposphere intensify with increasing height, as is to be expected from temperature profiles following moist adiabats to first order. Warmest and coldest temperatures are produced by the TTE and $2v_{\text{ice}}$ experiments, respectively. The 2-mom experiment stands out due to

a positive temperature anomaly that is limited to the region between 1 km to 3 km, where the largest negative $\mathcal{R}$ anomaly is found. This points to a shallower trade inversion in the 2-mom experiment. This could be indicative of an earlier onset of precipitation in the 2-mom experiment, resulting in clouds growing less deep (Stevens and Seifert, 2008).

Based on a simple analytical model Romps, 2014 showed that in radiative-convective equilibrium $\mathcal{R}$ should be an invariant function of temperature as the atmosphere warms. An obvious question is therefore if the changes in our sensitivity experiments are explained by an upward or downward shift of the $\mathcal{R}$ profile following an increase or decrease in temperature, respectively. This would mean that in experiments with a warmer troposphere $\mathcal{R}$ should increase in the lower and mid troposphere, where $\mathcal{R}$ decreases with height, and $\mathcal{R}$ should decrease in the upper troposphere, where $\mathcal{R}$ increases with height. While the TTE and 2-mom runs show a corresponding pattern in their $\mathcal{R}$ changes, the temperature differences between the experiments is by far not large enough to explain the $\mathcal{R}$ differences. This is evident when $\mathcal{R}$ is plotted as a function of temperature (not shown). We therefore conclude that the differences in $\mathcal{R}$ are not explained by a vertical shift following isotherms.

In summary, our experiments suggest that a large part of the $\mathcal{R}$ spread across today's GSRMs is can be explained by different formulations of small-scale mixing and cloud microphysical processes. At least in the limited number of experiments we performed, microphysical choices particularly impact $\mathcal{R}$ in a rather narrow altitude region associated with shallow convection, whereas the choice of the turbulence scheme affects $\mathcal{R}$ in a broader mid-tropospheric layer.

In the following we focus on $\mathcal{R}$ differences in the mid troposphere (4 km to 8 km, indicated by the gray lines in Figure B.19). Although mid-tropospheric $\mathcal{R}$ differences are, similar as in the DYAMOND ensemble, not particularly large, Lang et al., 2021 showed that $\mathcal{R}$ differences in this region are particularly important for differences in OLR.

B.4 LAGRANGIAN RECONSTRUCTIONS OF RELATIVE HUMIDITY

B.4.1 *Reconstructions based on the last-saturation model*

To obtain a better understanding of the physical mechanisms behind the humidity changes produced in our experiments we use a last-saturation framework based on back-trajectories. For this analysis we focus on the altitude region between 4 km and 8 km, where $\mathcal{R}$ differences in the DYAMOND ensemble were shown to have a comparably large effect on the clear-sky radiation budget (Lang et al., 2021). A main goal is to understand to what extent the changes in $\mathcal{R}$ are explained by changes in the properties of the source regions of air parcels, i.e. the points of last-saturation/condensation, and by changes in moisture sources and sinks

during subsequent advection.

To investigate this we perform Lagrangian reconstructions of $\mathcal{R}$ for the ICON experiments described in Section B.2. The reconstruction for each experiment is performed in two different ways. The first one is an implementation of the last-saturation paradigm similar to earlier studies (e.g. Dessler and Sherwood, 2000; Pierrehumbert and Roca, 1998; Sherwood, 1996), although the latter were based on much coarser wind fields from GCMs or reanalysis data. The underlying assumption is that specific humidity q is conserved after the last-condensation event. Hence, the specific humidity at a given target point q_t equals the specific humidity the respective parcel had when it last experienced condensation q_{lc}. $\mathcal{R}$ at the target point is then equal to

$$\mathcal{R}_{lc} = \frac{q_{lc}}{q_t^*}, \tag{11}$$

where q_t^* denotes the saturation specific humidity at the target point. q_{lc} should generally equal its saturation value q_{ls}^* (though supersaturation can occur with respect to ice), so that Equation 11 can be written as

$$\mathcal{R}_{lc} \approx \frac{q_{ls}^*}{q_t^*} = \frac{e^*(T_{lc})}{e^*(T_t)} \frac{p_t}{p_{lc}}, \tag{12}$$

where e^* is the saturation water vapour pressure, T_{lc} and T_t are the temperatures of the last-condensation point and the target point, respectively, and p_{lc} and p_t are the corresponding air pressures. Thus, if the last-saturation reconstruction captures the humidity changes in the ICON experiments, this means that they are explained by temperature and pressure changes between the source and target regions.

For the reconstructions we use the actual q_{lc} rather than q_{ls}^*, i.e. Equation 11 rather than Equation 12, since $\mathcal{R}$ is not always exactly 100% at the instant of last-condensation (see Section B.4.3). This slightly improves our reconstructions, but our main conclusions do not depend on whether or not $q_{lc} = q_{ls}^*$ is assumed for the last-saturation events. The terms last-condensation and last-saturation are used interchangeably in the following.

For the second reconstruction of $\mathcal{R}$ moisture sources and sinks s, which can change a parcel's water vapour content during its advection after the last-condensation event, are added:

$$\mathcal{R}_{lc+s} = \frac{q_{lc} + s}{q_t^*}. \tag{13}$$

s includes evaporation of hydrometeors that are transported with or sediment through an air parcel, as well as turbulent mixing. These processes are represented by the parameterizations of microphysics and turbulence in the ICON model. As we will show in Section B.4.6, the inclusion of these sources and sinks brings the reconstructed $\mathcal{R}$ closer to the ICON-simulated $\mathcal{R}$ (subsequently denoted by

$\mathcal{R}_{\mathrm{ICON}}$).

Using the reconstructions, the change in $\mathcal{R}$ between a sensitivity experiment and the control experiment can be decomposed into three contributions:

$$\Delta\mathcal{R}_{\mathrm{ICON}} = \Delta\mathcal{R}_{\mathrm{lc}} + \Delta(\mathcal{R}_{\mathrm{lc+s}} - \mathcal{R}_{\mathrm{lc}}) + \Delta r. \tag{14}$$

The first term on the right hand side represents changes in source and/or target region pressure and temperature. The second term denotes the effect of changes in parameterized moisture sources and sinks acting during advection to the target region. The residual r is the difference $\mathcal{R}_{\mathrm{ICON}} - \mathcal{R}_{\mathrm{lc+s}}$. It results from shortcomings in the reconstruction method (Sections B.4.2 to B.4.6), but also from the fact that the Lagrangian reconstruction does not include any numerical diffusion, as opposed to the Eulerian advection scheme in ICON. Hence, the Δr term includes changes in numerical diffusion, which might be important in the experiments with changed model resolution but is not captured by the Lagrangian reconstruction.

The methods used to determine the points of last-condensation and the moisture sources and sinks along back-trajectories are described in the following.

B.4.2 *Back-trajectories*

Back-trajectories are calculated offline using the ICON version of the trajectory tool LAGRANTO version 2.0 (Sprenger and Wernli, 2015; Wernli and Davies, 1997). An ensemble of 150,000 back-trajectories is released once per day at 12 UTC from randomly selected points in the tropics (30°S to 30°N) between 4 km and 8 km height. In the following we will refer to this region as the target region.

Comparing the $\mathcal{R}$ distribution of the 150,000 trajectory starting points to the one obtained from the full field showed that the sampling error is small compared to the $\mathcal{R}$ differences between the model experiments. By starting the trajectories at 12 UTC only, depending on longitude we sample at different local times and thus capture different phases of the diurnal cycle of free-tropospheric humidity. A comparison showed that when sampling at 0 UTC, the moistest tropical regions appear moister by about 2% than when sampling at 12 UTC. This is likely a signature of the diurnal cycle of global precipitation, which was highlighted by Stevens et al., 2019. The effect of the sampling on the humidity differences between two experiments is small because the effect of the diurnal cycle is similar in each experiment. As our main interest is in the differences between experiments we conclude that starting trajectories once per day is sufficient.

Trajectories are integrated backwards in time for 15 days based on 1-hourly instantaneous 3D model wind fields. Out of a total of 45 simulated days, due to the 15-day lead time for the back-trajectories and the omission of the first five simulated days due to model spinup, a 25-day period remains for the Lagrangian reconstructions.

Given that the trajectory calculations are based on hourly model wind fields, and that the transport algorithms we use neither share the same numerical methods used by the ICON model nor are performed on the same grid, individual trajectories are not accurate, in the sense that they do not necessarily follow the exact paths they would follow if they were calculated online during model integration (Miltenberger et al., 2013). However, from a large ensemble of back-trajectories it is possible to infer the statistical properties of the points of last condensation and subsequent moisture sources and sinks, as we will show in the following.

B.4.3 *Last-condensation events*

We define the point of last condensation to be the first point along a back-trajectory, for which the local moisture tendency from the microphysics parameterization $(\frac{\mathrm{d}q}{\mathrm{d}t})_{\mathrm{mic}}$ takes on a negative value, i.e. as the point at which condensation last occurred. We decided for this definition rather than using a threshold value on relative humidity, because the critical relative humidity for condensation in ICON can exceed 100% with respect to ice. As a result of the spatial interpolation of the model fields, both the interpolation from the native ICON grid to a latitude-longitude grid and the interpolation from the latitude-longitude grid onto the trajectory positions performed by LAGRANTO, there are points where $(\frac{\mathrm{d}q}{\mathrm{d}t})_{\mathrm{mic}} < 0$ (and are therefore detected as condensation points), but the local relative humidity is significantly smaller than 100%. We therefore introduce the additional condition that the local relative humidity must be higher than 80%. If this condition is not met, the search for a last-condensation event is continued backwards along the trajectory.

Last-condensation events identified by this method are subject to different uncertainties. Condensation events will be missed if they occur in between the 1-hourly model output time step, which our trajectories are calculated on. We expect this to introduce a dry bias in the reconstructed $\mathcal{R}$, since on average the identified last-condensation events occur too far in the past and therefore at too cold temperatures, assuming that most air parcels undergo subsidence on their way to the target region. Furthermore, the last-condensation events we determine are restricted to the 15-day period covered by the back-trajectories, so events occurring further in the past are not detected. We do not find a last-condensation event within 15 days for 7% of the trajectories. This is expected to introduce a moist bias in the reconstructed $\mathcal{R}$, assuming that the condensation events further back in time would occur at higher altitudes and therefore colder temperatures than the trajectory end points.

B.4.4 *Moisture sources and sinks from parameterized processes*

To estimate the magnitude of moisture sources and sinks S (Equation 13), along each trajectory we sum up the local tendencies of q from the microphysics and turbulence parameterizations $(\frac{\mathrm{d}q}{\mathrm{d}t})_{\mathrm{mic}}$ and $(\frac{\mathrm{d}q}{\mathrm{d}t})_{\mathrm{turb}}$, respectively, between the time of last condensation t_{lc} and the target point ($t = 0$):

$$s = \sum_{t=0}^{t_{\mathrm{lc}}} ((\frac{\mathrm{d}q}{\mathrm{d}t})_{\mathrm{mic,t}} + (\frac{\mathrm{d}q}{\mathrm{d}t})_{\mathrm{turb,t}})\Delta t, \tag{15}$$

where $\Delta t = 1\,\mathrm{h}$ is the model output interval. The moisture tendency from the turbulence scheme $(\frac{\mathrm{d}q}{\mathrm{d}t})_{\mathrm{turb}}$ output by ICON only includes the contribution from vertical mixing, although the Smagorinsky turbulence scheme also performs horizontal mixing. Including the contribution from horizontal mixing for one of the ICON experiments showed it to be negligible compared to the effect of vertical mixing.

B.4.5 *Spatial averaging*

Figures B.20a and B.20b show the (randomly chosen) start positions of back-trajectories for an exemplary simulation time step on a map. Each dot corresponds to one start position, colored by the ICON-simulated relative humidity ($\mathcal{R}_{\mathrm{ICON}}$) and reconstructed relative humidity ($\mathcal{R}_{\mathrm{lc+s}}$), respectively, for the respective position. Target regions for which $\mathcal{R}_{\mathrm{ICON}}$ takes on intermediate values show up as a mixture of very high and very low values in $\mathcal{R}_{\mathrm{lc+s}}$. This is likely due to the fact that gradients and extremes in $\mathcal{R}_{\mathrm{ICON}}$ are smoothed out due to the limited resolution of the ICON model. While each value of $\mathcal{R}_{\mathrm{ICON}}$ in Figure B.20a represents a grid-cell average, values of $\mathcal{R}_{\mathrm{lc+s}}$ in Figure B.20b represent structures (or "filaments") on smaller scales, which are not resolved on the ICON grid. To smooth the reconstructed fields the sampling would need to be improved by increasing the number of trajectories per ICON grid cell and averaging over them. Another source of noise in the reconstructed $\mathcal{R}$ are inaccuracies in the trajectories, which result from the coarse (1-hourly) temporal resolution and spatial interpolation of the input data (see Section B.4.2). These inaccuracies can result in last-condensation points being spatially displaced from their true position.

To minimize sampling biases and to make our analysis framework more commensurate with the information content in the input data we coarsen our analysis region by averaging all results within boxes that span an area of $2^\circ \times 2^\circ$ in the horizontal and the complete altitude range between 4 km and 8 km in the vertical. These boxes will be referred to as target boxes in the following. We predict the horizontally and vertically averaged relative humidity in each target box as the mean of $\mathcal{R}_{\mathrm{lc}}$, respectively $\mathcal{R}_{\mathrm{lc+s}}$, of all back-trajectories released from within the box. As shown in Figure B.20c and B.20d, there is good agreement between the spatially averaged $\mathcal{R}_{\mathrm{ICON}}$ and $\mathcal{R}_{\mathrm{lc+s}}$, though the reconstructed field is still a bit noisier.

For some trajectories, the Lagrangian reconstruction yields extreme, unphysically high values of $\mathcal{R}$. In these cases the last-condensation event occurred at higher temperatures than that of the target point, so the air parcels have ascended after the last-condensation event. The ascent and associated cooling would not be possible without further condensation, which would keep the air parcel's relative humidity at around 100%. However, due to the shortcomings in our method described in Sections B.4.2 and B.4.3, these further condensation events are missed and an extremely high value of $\mathcal{R}$ is predicted. We remove these cases prior to the spatial averaging by discarding trajectories for which $\mathcal{R}_{\mathrm{lc+s}}$ is more than 10% higher than the maximum of $\mathcal{R}_{\mathrm{ICON}}$, which is about 130% in the control experiment. This is the case for 5% of all trajectories for which a last-condensation event was determined.

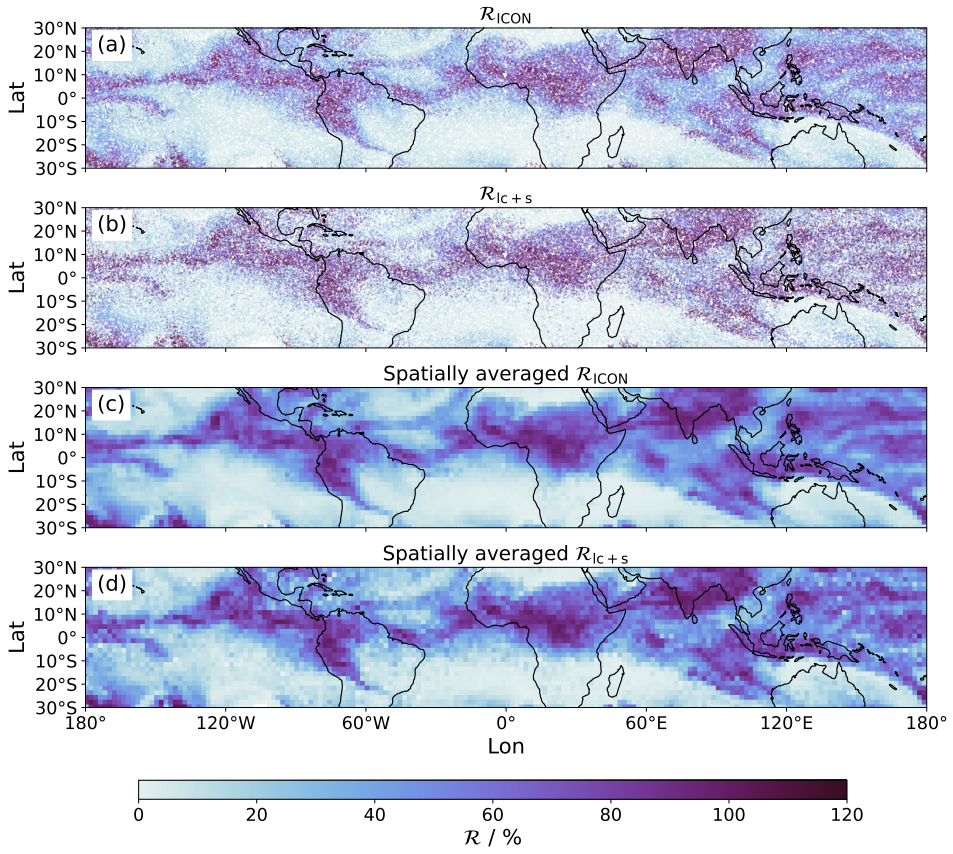


Figure B.20: Illustration of spatial averaging performed to reduce noise in the reconstructed relative humidity field for an exemplary time step (17 July 2021, 12Z). Scatterplots of (a) ICON-simulated relative humidity ($\mathcal{R}_{\mathrm{ICON}}$) and (b) reconstructed relative humidity $\mathcal{R}_{\mathrm{lc+s}}$ at the start positions of back-trajectories. Spatially averaged (c) $\mathcal{R}_{\mathrm{ICON}}$ and (d) $\mathcal{R}_{\mathrm{lc+s}}$ over $2^\circ \times 2^\circ$ boxes.

B.4.6 *Reconstructed relative humidity*

To evaluate the methods described above, we examine how well $\mathcal{R}_{\text{ICON}}$ is reproduced by Equations 11 and 13 in our control experiment. The distribution of $\mathcal{R}_{\text{ICON}}$ is bimodal with a prominent peak at values below 20% (Figure B.21). Such a bimodal distribution is well known from observations (e.g. Ryoo et al., 2009; Zhang et al., 2003) and has been attributed to the rapid drying by radiative subsidence; after being moistened by upward transport, air parcels dry out rapidly and spend a short time at intermediate humidity (Mapes, 2001).

Both kinds of Lagrangian reconstructions reproduce the ICON-simulated $\mathcal{R}_{\text{ICON}}$ well (Figure B.21). While the distribution of $\mathcal{R}_{\text{lc}}$ is shifted to lower values compared to $\mathcal{R}_{\text{ICON}}$, the distribution of $\mathcal{R}_{\text{lc+s}}$ is closer to, but shifted to slightly higher values than $\mathcal{R}_{\text{ICON}}$. The improvement of the reconstruction by including moisture sources and sinks is encouraging, as this would be expected if the approach was working as intended. The fact that the inclusion of moisture sources and sinks from the parameterizations increases the predicted relative humidity is not surprising. Per definition, microphysical processes can only increase an air parcel's q after the point of last condensation. Turbulent mixing can generally either increase or decrease q. However, vertical mixing, which dominates along our trajectories (see Section B.4.4), primarily moistens air parcels that subside through the free troposphere due to a down gradient moisture flux and the exponential decrease of q with height. Why $\mathcal{R}_{\text{lc+s}}$ tends to overestimate $\mathcal{R}_{\text{ICON}}$ is less clear and likely reflects uncertainties in our method and/or the fact that the Lagrangian reconstruction does not incorporate numerical diffusion. However, the aim of the Lagrangian reconstruction in this study is not to obtain a perfect reproduction of $\mathcal{R}_{\text{ICON}}$, but rather to explain humidity differences between different ICON experiments. As we will show in Section B.5.2, this is possible despite some small deviations of the $\mathcal{R}_{\text{lc+s}}$ distribution to the $\mathcal{R}_{\text{ICON}}$ distribution.

B.4.7 $\mathcal{R}$*-space*

To distinguish between different tropical humidity regimes, we divide the target boxes and the corresponding back-trajectories into ten equal-sized bins of $\mathcal{R}_{\text{ICON}}$. The driest bins in this "$\mathcal{R}$-space" correspond to the (sub-)tropical subsidence regions, whereas the moistest bins correspond to deep convective regions in the Intertropical Convergence Zone (ITCZ) and the Indo-Pacific Warm Pool. In our experiments, which are performed for a period in northern-hemisphere summer, the regions of highest $\mathcal{R}$ are centered around about 10° N and the driest regions are concentrated south of the equator, where the subsiding branch of the strong cross-equatorial Hadley cell is located (Figure B.22a). Regions of intermediate $\mathcal{R}$ are more widely distributed across the tropics, with a larger proportion located north of the equator.

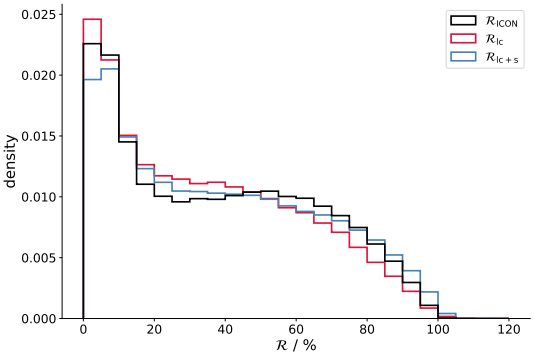


Figure B.21: ICON-simulated and reconstructed relative humidity distributions in the control experiment. Probability density of tropical relative humidity simulated by the ICON model ($\mathcal{R}_{\mathrm{ICON}}$, black) as well as from Lagrangian reconstructions based on the plain last-saturation model ($\mathcal{R}_{\mathrm{lc}}$, red) and taking into account moisture sources and sinks from parameterized processes ($\mathcal{R}_{\mathrm{lc+s}}$, blue). Histograms are based on $2^\circ \times 2^\circ$ spatially averaged relative humidity (see text for details).

The back-trajectories demonstrate how the origins of air parcels differ between regions of low and high $\mathcal{R}$. For the driest target regions south of the equator, last condensation occurs in two different regions remote from the target region: on the southern edge of the tropical deep convective regimes close to the equator, and in the sub- and extra-tropics (Figure B.22b). Towards regions of higher $\mathcal{R}$, the fraction of air parcels originating from within the tropics increases (Figure B.22c). Air parcels arriving in the driest regimes have on average travelled for about one week since last condensation (Figure B.22c), which is consistent with the time periods found by Cau et al., 2007 based on reanalysis fields. These air parcels have subsided from high altitudes, as evident from low last-condensation temperatures of about 220 K. The large difference between source and target temperature causes the extremely low target $\mathcal{R}$ of these parcels (Equation 12). In summary, regions of low $\mathcal{R}$ are characterized by source regions that are cold and remote. Towards regions of higher $\mathcal{R}$, last-condensation events occur closer to the target regions and at temperatures more similar to that of the target region (Figure B.22b,c). Air parcels arriving in the moistest target regions have travelled for less than a day since last condensation.

Figure B.23a shows mean and standard deviation of the reconstructed $\mathcal{R}_{\mathrm{lc}}$ and $\mathcal{R}_{\mathrm{lc+s}}$, respectively, plotted against mean $\mathcal{R}_{\mathrm{ICON}}$ for each bin in $\mathcal{R}$-space for the control experiment. The spread in the reconstructed $\mathcal{R}$ in each bin is comparable to the difference in $\mathcal{R}_{\mathrm{ICON}}$ between neighbouring bins, demonstrating that the Lagrangian reconstruction succeeds in predicting the $\mathcal{R}$ of a given target box. Again, it is evident that the plain last-saturation reconstruction underestimates $\mathcal{R}$, particularly in moist regimes, while the reconstruction with moisture sources and sinks slightly overestimates $\mathcal{R}$, particularly in dry regimes.

The difference between $\mathcal{R}_{\mathrm{lc+s}}$ and $\mathcal{R}_{\mathrm{lc}}$ provides an estimate of the effect of parameterized moisture sources on relative humidity. It increases from about 0.5%

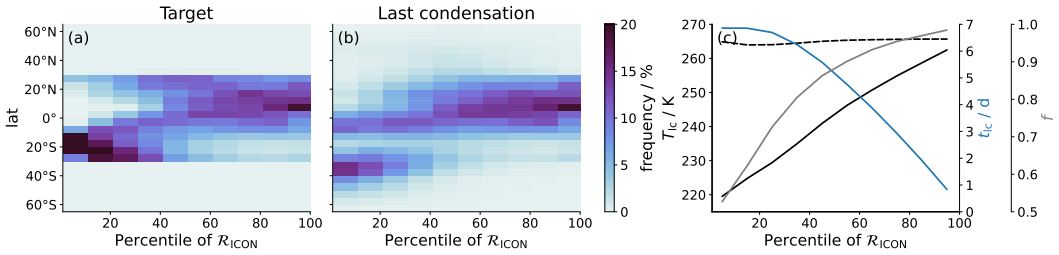


Figure B.22: Characteristics of target and source regions in the control experiment in $\mathcal{R}$-space. Histograms showing meridional distributions of (a) target regions and (b) last-condensation points for ten decile-bins of $\mathcal{R}_{\mathrm{ICON}}$. (c) Bin-averages of last-condensation temperature (T_{lc}, black solid) and time passed since last condensation (t_{lc}, blue), as well as fraction of last-condensation points located within the tropics, defined as 30° S to 30° N (f, gray). The temperature of the target region is denoted by the black dashed line.

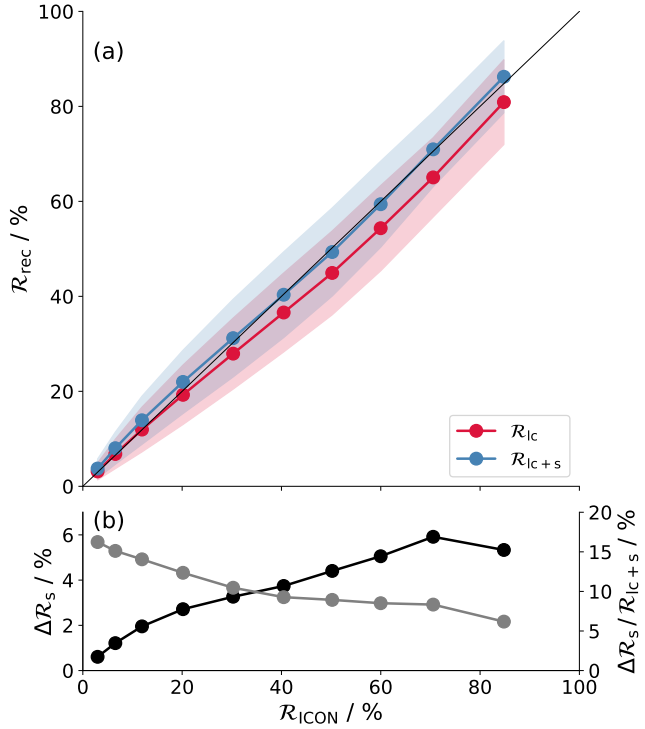


Figure B.23: ICON-simulated and reconstructed relative humidity $\mathcal{R}$ for the control experiment in $\mathcal{R}$-space. (a) Reconstructed $\mathcal{R}$ ($\mathcal{R}_{\mathrm{rec}}$) versus ICON-simulated $\mathcal{R}$ ($\mathcal{R}_{\mathrm{ICON}}$) for ten decile-bins of $\mathcal{R}_{\mathrm{ICON}}$. Points correspond to bin-mean values, the shading indicates ± one standard deviation of $\mathcal{R}_{\mathrm{rec}}$. Colours distinguish reconstructions based on the plain last-saturation model ($\mathcal{R}_{\mathrm{lc}}$, red) and taking into account moisture sources and sinks from parameterized processes ($\mathcal{R}_{\mathrm{lc+s}}$, blue). (b) The difference $\mathcal{R}_{\mathrm{lc+s}}-\mathcal{R}_{\mathrm{lc}}$ ($\Delta\mathcal{R}_{\mathrm{s}}$) in absolute units (black, left x-axis) and relative to $\mathcal{R}_{\mathrm{lc+s}}$ (gray, right x-axis).

in the driest decile to about 6% in the moistest decile (Figure B.23b). Although parcels that end up with low $\mathcal{R}$ also originate from moist regions, where microphysical processes and turbulent mixing are potentially active, they passed these regions at much colder temperatures (cf. Figure B.22c), at which water

vapor concentrations (and hence also sources) are small. Therefore, the effect from parameterized moisture sources on $\mathcal{R}$ increases from dry to moist regions when it is measured in absolute units. When the change in $\mathcal{R}$ from parameterized sources is measured relative to the final (reconstructed) value of $\mathcal{R}$ it decreases from about 15% in the driest decile to about 5% in the moistest decile. This reflects that the probability to encounter moisture sources is enhanced for parcels that end up with low $\mathcal{R}$, because they have been transported over a longer time since last condensation (cf. Figure B.22c). In general, the difference between $\mathcal{R}_{\mathrm{lc+s}}$ and $\mathcal{R}_{\mathrm{lc}}$ is small compared to the range of $\mathcal{R}$ values occurring throughout the tropics. This is in line with many earlier studies, which concluded that moisture sources and sinks are not relevant for explaining spatial variations of tropical $\mathcal{R}$ (e.g. Dessler and Sherwood, 2000; Sherwood, 1996), corroborating the general validity of the last-saturation paradigm. Nevertheless, they might be relevant for explaining more subtle $\mathcal{R}$ differences between model experiments. This will be examined in the course of this study.

B.5 MECHANISMS CONTROLLING MID-TROPOSPHERIC RELATIVE HUMIDITY DIFFERENCES

B.5.1 *Changes in mid-tropospheric relative humidity*

The representation of mid-tropospheric $\mathcal{R}$ differences in $\mathcal{R}$-space (Figure B.24a) shows that for most experiments changes in $\mathcal{R}$ are larger in moist than in dry regions. Therefore, differences in tropical mean $\mathcal{R}$ (Figure B.19) mainly reflect differences in the moist regions. A similar behaviour was also found for mid-tropospheric humidity differences among the DYAMOND models (Lang et al., 2021). The robustness of $\mathcal{R}$ in dry regions is related to their cold source temperatures, which will be discussed in more detail in Section B.5.3.

As already evident from the tropical mean $\mathcal{R}$ profiles, mid-trospheric $\mathcal{R}$ changes are largest in the experiment with the TTE turbulence scheme. The representation in $\mathcal{R}$-space shows that $\mathcal{R}$ increases throughout the tropics, but the strongest increase (about 10%) occurs in rather moist regimes around the 80th percentile of $\mathcal{R}$. In comparison, the sensitivity of mid-tropospheric $\mathcal{R}$ to changes in the microphysics (2-mom and $2v_{\mathrm{ice}}$) is weaker and limited to regions of intermediate and high $\mathcal{R}$. The experiment with halved vertical resolution ($2\Delta z$) is the only one in which changes in $\mathcal{R}$ are larger in dry than in moist regimes. The increase in mid-tropospheric $\mathcal{R}$ in the experiment with doubled horizontal resolution ($\Delta x/2$) is concentrated in moist regimes.

Internal variability, which we estimate from the difference between the two control experiments, is larger in dry than in moist regions. This may be expected given that the source regions of dry air are remote (Figure B.22) and therefore strongly influenced by the large-scale circulation, which varies on timescales that are longer than our simulation period. While in the moist regions (and therefore

also in the tropical mean) changes in $\mathcal{R}$ are larger than the estimated internal variability in all sensitivity experiments, in the dry regions this is only the case for the TTE and $2\Delta z$ experiments. Thus, the $\mathcal{R}$ differences we find in dry regions are strongly coloured by internal variability and systematic differences could only be quantified based on longer experiments. This should be kept in mind for the discussions in the following.

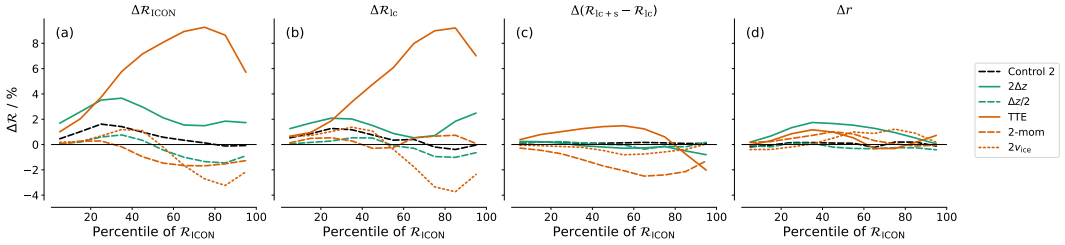


Figure B.24: ICON-simulated and reconstructed changes in mid-tropospheric $\mathcal{R}$ in the sensitivity experiments displayed in $\mathcal{R}$-space. (a) Changes in ICON-simulated $\mathcal{R}$ compared to the control experiment ($\Delta\mathcal{R}_{\mathrm{ICON}}$). (b) Changes in $\mathcal{R}$ reconstructed by a plain last-saturation model ($\Delta\mathcal{R}_{\mathrm{lc}}$) and (c) changes in the effect of moisture sources and sinks after last condensation ($\Delta(\mathcal{R}_{\mathrm{lc+s}} - \mathcal{R}_{\mathrm{lc}})$). (d) Changes in the residual (Δr), i.e. in the difference between ICON-simulated and reconstructed $\mathcal{R}$. The sum of the terms shown in (b) to (d) yields the ICON-simulated $\mathcal{R}$ changes shown in (a). Lagrangian reconstructions were not performed for the $\Delta x/2$ experiment (see text for explanation).

B.5.2 *Changes in source and target regions vs. changes during advection*

The two types of Lagrangian reconstructions (Equations 11 and 13) are used to shed light on the physical processes behind the $\mathcal{R}$ changes in the sensitivity experiments. The reconstructions were performed for all experiments except the $\Delta x/2$ experiment for reasons of limited resources as the doubled horizontal resolution increases the model output by a factor of four. Additionally, to obtain the same accuracy of trajectories as for the control experiment the timestep for the trajectory calculation would need to be halved. In total, the required model output for the $\Delta x/2$ experiment would increase by a factor of 8 and the trajectory calculations would get correspondingly expensive.

For most experiments the $\mathcal{R}$ differences that were reconstructed based on the plain last-saturation model ($\Delta\mathcal{R}_{\mathrm{lc}}$, Figure B.24b) explain a large part of the actual differences ($\Delta\mathcal{R}_{\mathrm{ICON}}$, Figure B.24a), whereas the effect from changes in parameterized processes given by $\Delta\mathcal{R}_{\mathrm{lc+s}} - \Delta\mathcal{R}_{\mathrm{lc}}$ is small (Figure B.24c). This means that the $\mathcal{R}$ changes must be mainly caused by changes in the source and/or target temperature (see also Section B.5.3), whereas changes in moisture sources and sinks that affect an air parcel's water vapor content on its way to the target region are of minor importance. Most importantly, different from what one might expect, the strong mid-tropospheric moistening in the TTE experiment is not a direct consequence of enhanced vertical turbulent mixing that moistens

air parcels as they are transported from source to target regions. Instead, it must be explained by changes in the properties of source and/or target regions themselves, which we will investigate further in later sections. Similarly, one might expect that the moistening in the $2\Delta z$ experiment with coarser vertical resolution results from enhanced numerical diffusion during vertical advection after last condensation. However, the moistening is at least partly reproduced by the Lagrangian reconstructions, which do not account for changes in numerical diffusion after last condensation. Having said this, the reconstructions do not fully capture the strong moistening of dry regions, which is also evident from the positive residual term (Figure B.24d). Hence, a part of the moistening might well be explained by enhanced numerical diffusion on the pathway from the source to the target point.

From the fact that the last-saturation model successfully reproduces the $\mathcal{R}$ changes between experiments, one could also conclude that they are caused by changes in the resolved circulation and the temperature structure. This is true under the assumption that the location (and hence temperature) of last-condensation points only depends on the resolved circulation and temperature structure. However, as we will explain in Section B.5.5, this assumption does not always hold.

There are exceptions, where changes in parameterized moisture sources and sinks after last condensation do play a role in changing $\mathcal{R}$. As one would expect, this mainly concerns the experiments with changes in the parameterizations of turbulence and microphysics. In the TTE experiment, turbulent moistening during advection is enhanced for dry and intermediate regimes and reduced for moist regimes. Overall, the contribution from the changing moisture sources to the total $\mathcal{R}$ change is small. The (rather weak) drying of the mid troposphere in the 2-mom experiment is mainly due to a reduction in moisture sources (Figure B.24c), while the plain last-saturation reconstruction predicts almost no change (Figure B.24b). Hence, the drying is caused by reduced evaporation of cloud condensate or precipitation. However, additional trajectory calculations showed that the stronger reduction in $\mathcal{R}$ in the layer between 1 km and 3 km in the 2-mom experiment (Figure B.19) is to a large extent captured by the plain last-saturation model. The ratio of air parcels that have subsided from the free troposphere since last condensation to air parcels that have very recently experienced saturation during ascent increases in the 2-mom experiment, indicating that the microphysical perturbation also affects the resolved transport associated with shallow convection. This would be consistent with the microphysics limiting the depth of shallow convection as mentioned in Section B.3.

The Δr term includes any changes in $\mathcal{R}_{\mathrm{ICON}}$ that are not explained by either of the two Lagrangian reconstructions (with or without moisture sources along the trajectory). As explained above, the positive Δr in the $2\Delta z$ experiment might result from an increase in numerical diffusion, which is not captured by the Lagrangian reconstruction. However, there are also a positive, albeit smaller Δr for the TTE,

2-mom and $2v_{\text{ice}}$ experiments, for which we do not expect changes in numerical diffusion.

In summary, the $\mathcal{R}$ changes in our experiments are largely explained by the last-saturation model, and only slightly modulated by changes in moisture sources after last condensation. In the $2\Delta z$ experiment the part of the $\mathcal{R}$ change that cannot be explained by either of the two mechanisms is likely related to changes in numerical diffusion.

B.5.3 *Changes in source temperature vs. changes in target temperature*

The fact that $\mathcal{R}$ differences are largely explained by the last-saturation model leaves changes in the saturation specific humidity in the source regions and in the target region as possible causes (Equation 12). With a linear expansion the relative humidity change predicted by the last-saturation model can be approximated as follows:

$$\Delta\mathcal{R}_{\text{lc}} \approx \frac{L_v}{R_v}\frac{\mathcal{R}_{\text{lc}}}{T_{\text{lc}}^2}\Delta T_{\text{lc}} - \frac{L_v}{R_v}\frac{\mathcal{R}_{\text{lc}}}{T_{\text{t}}^2}\Delta T_{\text{t}} = \Delta\mathcal{R}_{\text{s}} + \Delta\mathcal{R}_{\text{t}}, \tag{16}$$

where R_v is the gas constant of water vapor and L_v is the specific heat of vaporization of water. The first term $\Delta\mathcal{R}_{\text{s}}$ corresponds to the change in $\mathcal{R}_{\text{lc}}$ due to changes in source temperature, the second term $\Delta\mathcal{R}_{\text{t}}$ is the change in $\mathcal{R}_{\text{lc}}$ due to changes in target temperature. From Equation 12 there should be a third term representing changes in source pressure, which we found to be negligible compared to the temperature terms. Changes in target pressure do also not play a role since our target region is a fixed altitude region in all experiments.

$\Delta\mathcal{R}_{\text{s}}$ and $\Delta\mathcal{R}_{\text{t}}$ are shown in Figure B.25. Their sum is a good approximation of $\Delta\mathcal{R}_{\text{lc}}$ (not shown). The two terms tend to have opposite signs, indicating that an increase in last-condensation temperature, which increases $\mathcal{R}_{\text{lc}}$, is typically accompanied by an increase in the target temperature, which decreases $\mathcal{R}_{\text{lc}}$. However, $\Delta\mathcal{R}_{\text{s}}$ overcompensates $\Delta\mathcal{R}_{\text{t}}$ for all experiments. This is likely related to the fact that the source regions are generally located above the target regions (Figure B.22c) and temperature differences between experiments increase with height (Figure B.19d).

The overcompensation described above is also evident from the fact that changes in $\mathcal{R}$ (Figure B.24a) follow a similar pattern as changes in last-condensation temperature ΔT_{lc} (Figure B.26a). The 2-mom experiment is an exception, because its $\mathcal{R}$ change is controlled by a change in parameterized moisture sources after last condensation (Section B.5.2). As noted already in Section B.5.2, the magnitudes of $\mathcal{R}$ changes are damped towards dry regimes, although the magnitudes of ΔT_{lc} hardly change throughout $\mathcal{R}$-space. This is because the absolute temperature of the source regions T_{lc} increases from dry to moist regimes (Figure B.22c). Due to the non-linear dependence of e^* on T the same temperature change results in a smaller change in e^*_{lc} at lower temperatures than at higher temperatures, and

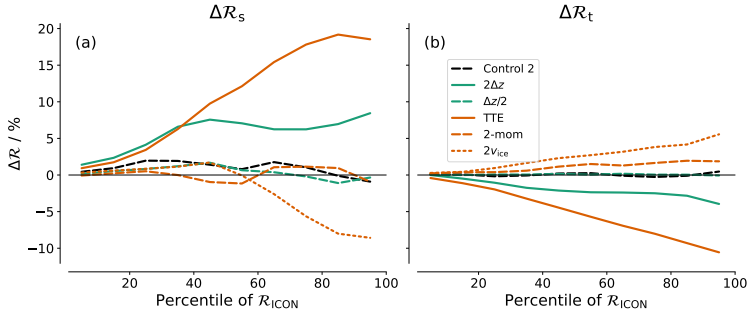


Figure B.25: Contributions from source and target temperature changes to changes in mid-tropospheric $\mathcal{R}$ in the sensitivity experiments shown in $\mathcal{R}$-space. (a) Contribution from change in last-condensation temperature ($\Delta\mathcal{R}_s$) and (b) contribution from change in target temperature ($\Delta\mathcal{R}_t$). The sum of two terms approximates the $\mathcal{R}$ changes that were reconstructed based on the last-saturation model ($\mathcal{R}_{lc}$ in Figure B.24b). Note the different in y-axis ranges in this figure and Figure B.24.

hence in a smaller change in $\mathcal{R}$. Thus, the robustness of $\mathcal{R}$ in dry regions is a consequence of the low water vapor concentrations in the cold source regions.

B.5.4 *Changes in tropical source regions vs. changes in extra-tropical source regions*

The source regions of tropical mid-tropospheric air lie both within the tropics (here defined as 30°S to 30°N) and in the extra-tropics (Figure B.22). Hence, changes in T_{lc} could result from changes in tropical last-condensation temperatures $T_{lc,trop}$, extra-tropical last-condensation temperatures $T_{lc,extra}$ or the share of tropical last-condensation points f:

$$\Delta T_{lc} \approx f\Delta T_{lc,trop} + (1-f)\Delta T_{lc,extra} + \Delta f(T_{lc,trop} - T_{lc,extra}) \tag{17}$$

In moist regimes, the changes in T_{lc} are dominated by changes in $T_{lc,trop}$ (Figure B.26b), whereas in the driest 40 percentiles changes in $T_{lc,trop}$ and $T_{lc,extra}$ are commensurately important (Figure B.26c). Note that the fraction of tropical last-condensation events f shapes the lines in Figure B.26 b and c. While the absolute changes in T_{lc} are similar for tropics and extra-tropics (not shown), extra-tropical changes do not affect the moist regions because f is close to 1 there (Figure B.22). Changes in f between experiments play a minor role in changing T_{lc} (Figure B.26d).

Internal variability (as measured by the Control 2 simulation) increases towards dry regions both for tropical and extra-tropical source regions (Figure B.26b,c). For the extra-tropics, changes in most sensitivity experiments are similar in magnitude and go in the same direction as in the Control 2 experiment, which may indicate that the control climate was an outlier with colder extra-tropical source temperatures. This explains why in the control experiment the driest regions have a lower $\mathcal{R}$ than in all the sensitivity experiments (Figure B.24). Thus, to a large extent, changes in $T_{lc,extra}$ in our sensitivity experiments can be explained

by, or at least not differentiated from, internal variability. This variability is likely caused by changes in the dynamic mechanisms that bring air to saturation in the extratropics and transport it to the tropics. The fact that the relative humidity of the dry regions is disproportionately affected by these changes emphasizes the important role of the exchange between extra-tropics and tropics in controlling the humidity of the dry regions, which has been highlighted in several studies (e.g. Cau et al., 2007; Roca et al., 2012; Villiger et al., 2022; Waugh, 2005). In particular, a change in these exchange mechanisms under warming represents a possible pathway for changing the relative humidity of the dry regions.

A change in $T_{\mathrm{lc,trop}}$ can generally result from a change in the tropical temperature profile and/or a change in the height distribution of last-condensation points. Additional analysis showed that both mechanisms are of similar importance in our experiments. Depending on the experiment they either counteract or reinforce each other. In the TTE experiment, for example, tropical temperature increases (see Figure B.19d) and last condensation occurs at lower altitudes on average. Both effects increase $T_{\mathrm{lc,trop}}$. In the $2v_{\mathrm{ice}}$ experiment, on the other hand, the two effects counteract; tropical temperature decreases, but last-saturation takes place at lower altitudes on average. This explains why the $\mathcal{R}$ change in the $2v_{\mathrm{ice}}$ experiment is relatively small despite the large temperature change (Figure B.19).

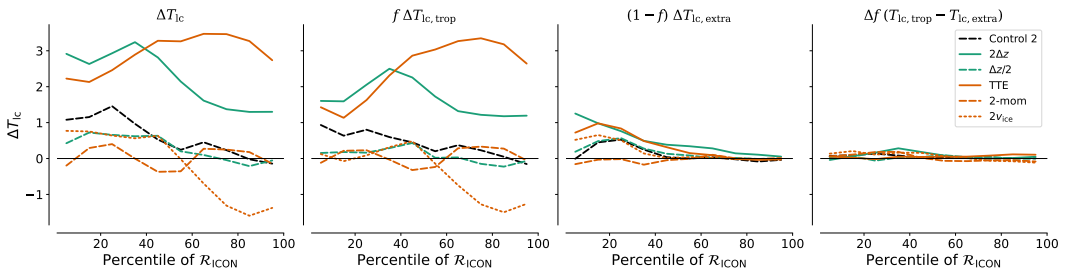


Figure B.26: Changes in last-condensation temperature T_{lc} in sensitivity experiments shown in $\mathcal{R}$-space. (a) Total change of T_{lc} compared to the control experiment, (b) contribution from changes in tropical last-condensation temperatures $T_{\mathrm{lc,trop}}$, (c) contribution from changes in extra-tropical last-condensation temperatures $T_{\mathrm{lc,extratrop}}$ and (d) contribution from changes in f, the share of tropical last-condensation events.

B.5.5 *Mechanisms behind the moistening in the TTE experiment*

Mid-tropospheric $\mathcal{R}$ increases most strongly in the experiment with the TTE turbulence parameterization. The analysis above has shown that this moistening is largely explained by an increase in the average temperature at last condensation. The full distribution of tropical last-condensation temperature $T_{\mathrm{lc,trop}}$ for the control and the TTE experiment are shown in Figure B.27. It is apparent that the distribution is bimodal in both experiments, implying that there are two distinct source regions for tropical mid-tropospheric air. The warm mode at around 265 K corresponds to “young" air parcels with high $\mathcal{R}$ that either experienced last

condensation very recently and have since subsided over only a short distance or are even saturated at the time considered. The cold mode at around 220 K represents “old" air parcels that have subsided from the upper troposphere, where deep convection detrains preferentially, and hence end up with a low $\mathcal{R}$ in the mid troposphere. In the TTE experiment the two modes stay at roughly the same temperature as in the control experiment, but the share of young air parcels increases at the expense of old air parcels. In line with that, snapshots of $\mathcal{R}$ and moisture tendencies from microphysics reveal that condensation occurs over a broader area of the tropical mid troposphere at any given time in the TTE experiment (not shown).

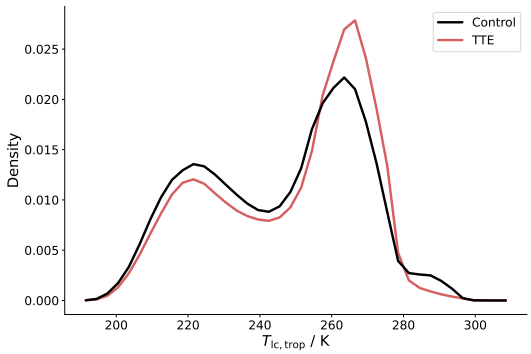


Figure B.27: Probability density distribution of last-condensation temperature T_{lc} for tropical last-condensation points in the control (black) and TTE (red) experiments.

One possible explanation for the broadening of saturated mid-tropospheric regions would be that convective updrafts cover a larger area. However, a corresponding analysis showed that this is not the case in the TTE experiment. The reason rather appears to be a strong turbulent mixing between lower and mid troposphere performed by the TTE scheme. Figure B.28a shows vertical profiles of the specific humidity tendencies produced by the turbulence scheme in the control and TTE experiments for an exemplary model output timestep. To distinguish between different tropical large-scale circulation regimes, profiles were averaged within five 20-percentile ranges of column-integrated water vapor. In the control experiment the Smagorinsky turbulence scheme only acts within the boundary layer throughout all circulation regimes; the air within the boundary layer is moistened by mixing water vapor upward from the surface. The TTE scheme behaves very differently. Most importantly, it performs a strong mixing between the lower and mid troposphere, particularly in the moist tropics, which manifests as a drying of the lower troposphere and a moistening of the mid troposphere. In other words, the TTE scheme unintentionally acts similar to a convective parameterization.

The mid-tropospheric moistening by turbulent mixing in the TTE experiment is accompanied by increased condensation, as evident from the specific humidity tendencies produced by the microphysics parameterization shown in Figure B.28b. The strong vertical mixing creates a moist background that favours condensation whenever air is displaced upward, such that condensation is not restricted to

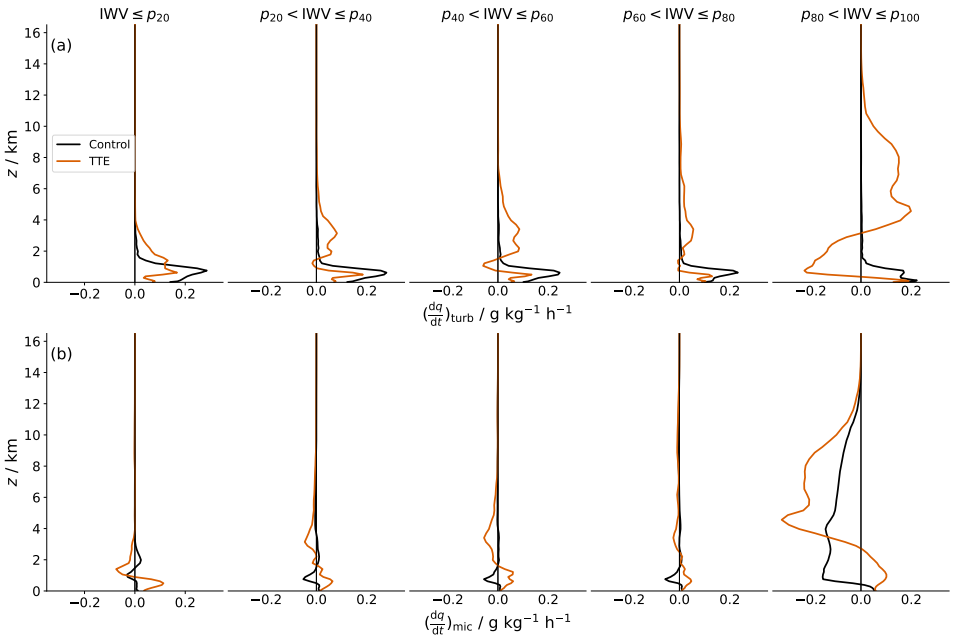


Figure B.28: Moisture tendencies from (a) turbulence and (b) microphysics parameterizations in the control (black) and TTE (red) experiment for an exemplary simulation time step (17 July 2021, 12Z). Each panel in (a) and (b) shows a vertical profile of specific humidity tendencies averaged over a 20-percentile range of column-integrated water vapor, sorted from dry profiles on the left to moist profiles on the right.

convective updrafts in the TTE experiment. This explains why the share of young air parcels with last condensation within the mid troposphere is increased.

It is worth revisiting Figure B.24c, which shows how the effect of parameterized moisture sources changed compared to the control simulation. Given that the turbulent moistening of the free troposphere is more intense in the TTE experiment, it may be surprising that for the moist percentiles the moistening from parameterized processes decreased compared to the control run. However, it can be understood as a consequence of the larger share of young air parcels in the moist percentiles, for which the time period available for moistening is reduced. This is also evident from Figure B.29, which shows the relative change in time since last condensation (t_{lc}) to the control experiment for all sensitivity experiments. In the TTE experiment, parcels arriving in the moistest percentile have on average been transported for a more than 40% shorter time since last condensation. For the other experiments changes in t_{lc} are within $\pm$ 10%.

While the last-saturation model technically explains the $\mathcal{R}$ increase in the TTE experiment, it does not do so for the reasons we expected. The original idea was that last-condensation points are determined by the resolved circulation and temperature structure. Thus, if the change in $\mathcal{R}$ is explained by the last-saturation model, it must be caused by changes in circulation and temperature, while changes in parameterized processes can only play a role if they affect these re-

solved properties. In the TTE experiment, however, condensation is not exclusively driven by resolved upward motions, but also by the strong parameterized vertical mixing of water vapor. Thus, in this case, parameterized moisture sources directly influence the location of the last-condensation events. Nevertheless, the fact that the last-saturation model succeeds in reproducing the $\mathcal{R}$ change still tells us that the change is driven by changes within the tropical source regions, i.e. the ITCZ and warm pool region, whereas changes in moisture sources during subsequent advection play a minor role.

The behavior of the TTE scheme is certainly unexpected and indicates that the scheme has not been sufficiently adapted to storm-resolving resolutions. Whether this type of one-dimensional scheme is appropriate for use at storm-resolving resolution is a question to be addressed in other studies. Having said that, the fact that even this extreme perturbation did not change $\mathcal{R}$ far beyond the inter-model spread in DYAMOND is promising. Many of the DYAMOND models used turbulence parameterizations that were not specifically adapted to storm-resolving resolution due to their early development stage. Hence, a better adaption of the schemes in future model versions might further reduce the spread in tropical $\mathcal{R}$.

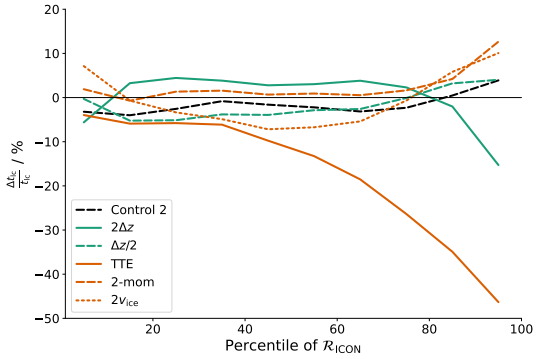


Figure B.29: Relative change in time since last condensation (t_{lc}) to the control experiment for all sensitivity experiments depicted in $\mathcal{R}$-space.

B.6 SUMMARY AND CONCLUSIONS

In this study our aim was to narrow down the model uncertainties that cause the remaining spread in tropical relative humidity $\mathcal{R}$ across GSRMs, as has been quantified in a recent study based on DYAMOND, the first model intercomparison initiative for GSRMs. To this end, we test the sensitivity of $\mathcal{R}$ to changes in model resolution and parameterizations in a series of six 45-day experiments with the ICON model in a storm-resolving configuration. The changes we apply to the model are inspired by differences among the DYAMOND models. They include changes in horizontal and vertical grid spacing, as well as in the parameterizations of microphysics and turbulence. We use a last-saturation model based on 3D backward trajectories to gain insight into the mechanisms behind the $\mathcal{R}$ changes

in the sensitivity experiments. This analysis is restricted to the mid troposphere.

The rather strong perturbations applied in our sensitivity experiments result in changes in tropical $\mathcal{R}$ that are of similar magnitude as the spread across the DYAMOND models. An earlier study had shown based on the DYAMOND ensemble that the $\mathcal{R}$ spread across GSRMs is reduced compared to classical GCMs with convective parameterizations. Our experiments support this finding by showing that even strong perturbations in GSRMs cannot reproduce the spread in $\mathcal{R}$ seen in models with convective parameterizations. Moreover, our experiments show that tropical $\mathcal{R}$ is rather robust to changes in model resolution within the general scale of GSRM resolutions. The three experiments with different vertical grid spacing (800 m, 400 m and 200 m in the free troposphere) show that $\mathcal{R}$ changes are modest as soon as a certain threshold vertical resolution is exceeded. The experiments with 5 km and 2.5 km horizontal grid spacing produce a very similar $\mathcal{R}$ distribution. While these results suggest that differences in model resolution do not contribute significantly to the current $\mathcal{R}$ spread across GSRMs, it does not exclude the possibility that reducing the horizontal grid spacing to much finer scales (on the order of 200 m) could make a difference, which needs to be tested in future experiments.

In our experiments, $\mathcal{R}$ changes more strongly in response to exchanging the microphysics and turbulence schemes, indicating that the model physics rather than resolution (at storm-resolving scales) are the major source of $\mathcal{R}$ spread across GSRMs. While microphysical changes affect $\mathcal{R}$ most strongly in the altitude layer associated with shallow clouds, exchanging the turbulence scheme changes $\mathcal{R}$ over a broad altitude region in the lower to mid troposphere. We could not test the extent to which the dynamical core, and choices it makes in how to solve the transport equations, systematically influences the distribution of source regions. However, the similarity of spread between our (parameterized) physics sensitivity studies, and the relatively modest effect of grid spacing lead us to believe that these effects are unlikely large.

Like the $\mathcal{R}$ differences between DYAMOND models, the $\mathcal{R}$ changes in our experiments are smallest in the dry subsidence regimes of the tropics. This is a consequence of the low water vapor concentrations in their cold source regions. However, since the sensitivity of OLR to changes in relative humidity is particularly high in dry background states (e.g. Spencer and Braswell, 1997), small $\mathcal{R}$ differences in the dry zones are nevertheless important from a radiative perspective (Lang et al., 2021). At the same time, this study highlights that understanding humidity differences between models is particularly challenging for the dry regions. The $\mathcal{R}$ of the dry regions is subject to larger internal variability on timescales of days to months, which storm-resolving simulations are currently limited to. This is because the source regions of dry air are located remotely (mainly on the edges of the inner-tropical deep-convective regimes and in the extra-tropics) and therefore depend on the large-scale circulation. Thus, while one simulated month is sufficient to quantify systematic $\mathcal{R}$ differences in moist

regions, longer simulations would increase our confidence in the sources of variability in the dry regions. Because changes in both tropical and extra-tropical origins need to be considered to understand $\mathcal{R}$ differences in dry regions (see also Cau et al., 2007; Roca et al., 2012), changes in the mechanisms of exchange between tropics and extra-tropics in a warmer climate represent an important pathway for changing the relative humidity of the dry regions, which would have important implications for the clear-sky climate feedback.

The mid-tropospheric $\mathcal{R}$ changes in our experiments, including the strong moistening in the experiment with the exchanged turbulence scheme, are largely captured by the last-saturation model. This means that most $\mathcal{R}$ changes are explained by changes in source temperature, i.e. the temperature at which air parcels typically experience last condensation, whereas changes in the moistening or drying by parameterized processes after last condensation play a minor role. This is even true when the parameterized moisture sources are modified directly, like in our microphysics and turbulence experiments. Overall, this study shows that the last-saturation model is not only successful in explaining variations in tropical $\mathcal{R}$ in the real atmosphere or a given model, as shown by many previous studies (e.g. Dessler and Sherwood, 2000; Pierrehumbert and Roca, 1998; Sherwood, 1996), but it can also be a helpful tool for explaining the causes of humidity differences between models. However, it has also become clear that last-saturation statistics can be directly affected by changes in parameterized moisture sources, e.g. by enhanced turbulent moistening. Therefore, if the last-saturation model explains a change in $\mathcal{R}$, it does not necessarily mean that it is due to changes in the resolved circulation or the temperature structure.

In our experiments the most substantial change in $\mathcal{R}$ was found in response to changing the turbulence parameterization from a Smagorinsky-type scheme to a total turbulent energy (TTE) scheme. The resulting increase in $\mathcal{R}$ was largest in the mid troposphere of moist regions. The reason appears to be that the TTE scheme produces a strong turbulent moistening of the mid troposphere in the inner, moist tropics. This moistening favours condensation, which is why from a last-saturation perspective the share of young air parcels with warm source temperatures increases in the TTE experiment. Thus, the $\mathcal{R}$ of the moist tropical regions, while less radiatively important than the dry regions, is disproportionally sensitive to vertical mixing processes that structure the humidity through their effect on the last-saturation temperatures, i.e. by increasing mid-level cloudiness, rather than their effect on the evolution of humidity since its last-saturation.

While the behavior of the TTE scheme is certainly unexpected and indicates that the scheme is poorly adapted to storm-resolving resolutions, the fact that even this extreme perturbation does not change $\mathcal{R}$ beyond the differences in the DYAMOND ensemble is very promising. Due to their early development stage, many of the DYAMOND models in fact used turbulence parameterizations that were not specifically adjusted to storm-resolving resolution. This nourishes hopes that

tropical relative humidity will become even more consistent across future model versions with better adapted schemes.

OPEN RESEARCH

The ICON model code is available on `https://mpimet.mpg.de/en/science/modeling-with-icon/code-availability`.

The simulation runscripts and the code producing the plots from post-processed model output and trajectories is available on Zenodo through `https://doi.org/10.5281/zenodo.7120534`.

ACKNOWLEDGEMENTS

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.

REFERENCES

Aemisegger, Franziska et al. (2021). "How Rossby wave breaking modulates the water cycle in the North Atlantic trade wind region." In: *Weather and Climate Dynamics* 2.1, pp. 281–309. DOI: 10.5194/wcd-2-281-2021.

Baldauf, Michael et al. (2011). "Operational Convective-Scale Numerical Weather Prediction with the COSMO Model: Description and Sensitivities." In: *Monthly Weather Review* 139.12, pp. 3887–3905. DOI: 10.1175/MWR-D-10-05013.1.

Bony, S. et al. (2006). "How Well Do We Understand and Evaluate Climate Change Feedback Processes?" In: *Journal of Climate* 19.15, pp. 3445–3482. DOI: 10.1175/JCLI3819.1.

Bourdin, S., L. Kluft, and B. Stevens (2021). "Dependence of Climate Sensitivity on the Given Distribution of Relative Humidity." In: *Geophysical Research Letters*. DOI: https://doi.org/10.1029/2021GL092462.

Bryan, George H., John C. Wyngaard, and J. Michael Fritsch (2003). "Resolution Requirements for the Simulation of Deep Moist Convection." In: *Monthly Weather Review* 131.10, pp. 2394–2416. DOI: 10.1175/1520-0493(2003)131<2394:RRFTSO>2.0.CO;2.

Cau, Piero, John Methven, and Brian Hoskins (2007). "Origins of Dry Air in the Tropics and Subtropics." In: *Journal of Climate* 20.12, pp. 2745–2759. DOI: 10.1175/JCLI4176.1.

Dessler, A. E. and S. C. Sherwood (2000). "Simulations of tropical upper tropospheric humidity." In: *Journal of Geophysical Research* 105.D15, pp. 20155–20163. DOI: 10.1029/2000JD900231.

Dipankar, Anurag et al. (2015). "Large eddy simulation using the general circulation model ICON." In: *Journal of Advances in Modeling Earth Systems* 7.3, pp. 963–986. DOI: 10.1002/2015MS000431.

ECMWF (2018). "IFS Documentation CY45R1." In: chap. Part IV : Physical processes. URL: https://www.ecmwf.int/node/18714.

Held, I. M. and K. M. Shell (2012). "Using Relative Humidity as a State Variable in Climate Feedback Analysis." In: *Journal of Climate* 25.8, pp. 2578–2582. DOI: 10.1175/JCLI-D-11-00721.1.

Held, I. M. and B. J. Soden (2000). "Water Vapour Feedback and Global Warming." In: *Annual Review of Energy and the Environment* 25.1, pp. 441–475. DOI: 10.1146/annurev.energy.25.1.441.

Heymsfield, Andrew J. and Leo J. Donner (1990). "A Scheme for Parameterizing Ice-Cloud Water Content in General Circulation Models." In: *Journal of the Atmospheric Sciences* 47.15, pp. 1865–1877. DOI: 10.1175/1520-0469(1990)047<1865:ASFPIC>2.0.CO;2.

Hohenegger, Cathy et al. (2022). "ICON-Sapphire: simulating the components of the Earth System and their interactions at kilometer and subkilometer scales." In: DOI: 10.5194/gmd-2022-171.

John, V. O. and B. J. Soden (2007). "Temperature and humidity biases in global climate models and their impact on climate feedbacks." In: *Geophysical Research Letters* 34.18, p. L18704. DOI: 10.1029/2007GL030429.

Kluft, Lukas et al. (2019). "Re-Examining the First Climate Models: Climate Sensitivity of a Modern Radiative–Convective Equilibrium Model." In: *Journal of Climate* 32.23, pp. 8111–8125. DOI: 10.1175/JCLI-D-18-0774.1.

Lang, Theresa et al. (2021). "Tropical Free-Tropospheric Humidity Differences and Their Effect on the Clear-Sky Radiation Budget in Global Storm-Resolving Models." In: *Journal of Advances in Modeling Earth Systems* 13.11. DOI: 10.1029/2021MS002514.

Lilly, D. K. (1962). "On the numerical simulation of buoyant convection." In: *Tellus* 14.2, pp. 148–172. DOI: 10.3402/tellusa.v14i2.9537.

— (1967). "The Representation of Small-Scale Turbulence in Numerical Simulation Experiments." In: *IBM Scientifc Computing Symposium on Environmental Sciences*. Ed. by H. H. Goldstine. Yorktown Heights, New York, pp. 195–210.

Mapes, B. E. (2001). "Water's two height scales: The moist adiabat and the radiative troposphere." In: *Quaterly Journal of the Royal Meteorological Society* 127.577, pp. 2353–2366. DOI: 10.1002/qj.49712757708.

Mauritsen, Thorsten et al. (2007). "A Total Turbulent Energy Closure Model for Neutrally and Stably Stratified Atmospheric Boundary Layers." In: *Journal of the Atmospheric Sciences* 64.2, pp. 645–655. DOI: 10.1175/2007jas2294.1.

Mauritsen, Thorsten et al. (2022). "Early development and tuning of a global coupled cloud resolving model, and its fast response to increasing CO2." In: DOI: 10.31223/X5T933.

McKim, Brett A., Nadir Jeevanjee, and Geoffrey K. Vallis (2021). "Joint Dependence of Longwave Feedback on Surface Temperature and Relative Humidity." In: *Geophysical Research Letters* 48.18. DOI: 10.1029/2021GL094074.

Miltenberger, A. K., S. Pfahl, and H. Wernli (2013). "An online trajectory module (version 1.0) for the nonhydrostatic numerical weather prediction model COSMO." In: *Geoscientific Model Development* 6.6, pp. 1989–2004. DOI: 10.5194/gmd-6-1989-2013.

Morrison, H., J. A. Curry, and V. I. Khvorostyanov (2005). "A New Double-Moment Microphysics Parameterization for Application in Cloud and Climate Models. Part I: Description." In: *Journal of the Atmospheric Sciences* 62.6, pp. 1665–1677. DOI: 10.1175/JAS3446.1.

Phillips, Vaughan T. J., Leo J. Donner, and Stephen T. Garner (2007). "Nucleation Processes in Deep Convection Simulated by a Cloud-System-Resolving Model with Double-Moment Bulk Microphysics." In: *Journal of the Atmospheric Sciences* 64.3, pp. 738–761. DOI: 10.1175/JAS3869.1.

Pierrehumbert, R. T. and R. Roca (1998). "Evidence for Control of Atlantic Subtropical Humidity by Large Scale Advection." In: *Geophysical Research Letters* 25.24, pp. 4537–4540. DOI: 10.1029/1998GL900203.

Pincus, Robert, Eli J. Mlawer, and Jennifer S. Delamere (2019). "Balancing Accuracy, Efficiency, and Flexibility in Radiation Calculations for Dynamical Models." In: *Journal of Advances in Modeling Earth Systems* 11.10, pp. 3074–3089. DOI: 10.1029/2019MS001621.

Po-Chedley, S. et al. (2019). "Climatology Explains Intermodel Spread in Tropical Upper Tropospheric Cloud and Relative Humidity Response to Greenhouse

Warming." In: *Geophysical Research Letters* 46.22, pp. 13399–13409. DOI: 10.1029/2019GL084786.

Raddatz, T. J. et al. (2007). "Will the tropical land biosphere dominate the climate–carbon cycle feedback during the twenty-first century?" In: *Climate Dynamics* 29.6, pp. 565–574. DOI: 10.1007/s00382-007-0247-8.

Roca, R. et al. (2012). "Tropical and Extra-Tropical Influences on the Distribution of Free Tropospheric Humidity Over the Intertropical Belt." In: *Surveys in Geophysics* 33, pp. 565–583. DOI: 10.1007/s10712-011-9169-4.

Romps, D. M. (2014). "An Analytical Model for Tropical Relative Humidity." In: *Journal of Climate* 27.19, pp. 7432–7449. DOI: 10.1175/JCLI-D-14-00255.1.

Ryoo, J.-M., T. Igusa, and D. W. Waugh (2009). "PDFs of Tropical Tropospheric Humidity: Measurements and Theory." In: *Journal of Climate* 22.12, pp. 3357–3373. DOI: 10.1175/2008JCLI2747.1.

Satoh, M. et al. (2019). "Global Cloud-Resolving Models." In: *Current Climate Change Reports* 5.3, pp. 172–184. DOI: 10.1007/s40641-019-00131-0.

Seifert, Axel and Klaus D. Beheng (2001). "A double-moment parameterization for simulating autoconversion, accretion and selfcollection." In: *Atmospheric Research* 59-60, pp. 265–281. DOI: 10.1016/S0169-8095(01)00126-0.

Sherwood, S. C. (1996). "Maintainance of the Free-Tropospheric Tropical Water Vapor Distribution. Part II: Simulation by Large-Scale Advection." In: *Journal of Climate* 9.11, pp. 2919–2934. DOI: 10.1175/1520-0442(1996)009<2919:MOTFTT>2.0.CO;2.

Sherwood, S. C. et al. (2010). "Tropospheric Water Vapor, Convection and Climate." In: *Reviews of Geophysics* 48, RG2001. DOI: 10.1029/2009RG000301.

Smagorinsky, J. (1963). "General Circulation Experiments with the Primitive Equations." In: *Monthly Weather Review* 91.3, pp. 99–164. DOI: 10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2.

Spencer, R. W. and W. D. Braswell (1997). "How Dry is the Tropical Free Troposphere? Implications for Global Warming Theory." In: *Bulletin of the American Meteorological Society* 78.6, pp. 1097–1106. DOI: 10.1175/1520-0477(1997)078<1097:hdittf>2.0.co;2.

Sprenger, M. and H. Wernli (2015). "The LAGRANTO Lagrangian analysis tool – version 2.0." In: *Geoscientific Model Development* 8.8, pp. 2569–2586. DOI: 10.5194/gmd-8-2569-2015.

Stevens, Bjorn et al. (2019). "DYAMOND: the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains." In: *Progress in Earth and Planetary Science* 6.61. DOI: 10.1186/s40645-019-0304-z.

Stevens, Björn and Axel Seifert (2008). "Understanding macrophysical outcomes of microphysical choices in simulations of shallow cumulus convection." In: *Journal of the Meteorological Society of Japan* 86A, pp. 143–162. DOI: 10.2151/jmsj.86A.143.

Vial, J., J.-L. Dufresne, and S. Bony (2013). "On the interpretation of inter-model spread in CMIP5 climate sensitivity estimates." In: *Climate Dynamics* 41.11-12, pp. 3339–3362. DOI: 10.1007/s00382-013-1725-9.

Villiger, Leonie et al. (2022). "Lagrangian formation pathways of moist anomalies in the trade-wind region during the dry season: two case studies from

EUREC&ampltsup&ampgt4&amplt/sup&ampgtA." In: *Weather and Climate Dynamics* 3.1, pp. 59–88. DOI: 10.5194/wcd-3-59-2022.

Waugh, Darryn W. (2005). "Impact of potential vorticity intrusions on subtropical upper tropospheric humidity." In: *Journal of Geophysical Research* 110.D11. DOI: 10.1029/2004JD005664.

Wernli, Heini and Huw C. Davies (1997). "A Lagrangian-based analysis of extratropical cyclones. I: The method and some applications." In: *Quarterly Journal of the Royal Meteorological Society* 123.538, pp. 467–489. DOI: 10.1002/qj.49712353811.

Zhang, C., B. E. Mapes, and B. J. Soden (2003). "Bimodality in tropical water vapour." In: *Quarterly Journal of the Royal Meteorological Society* 129.594, pp. 2847–2866. DOI: 10.1256/qj.02.166.

Zängl, Günther et al. (2015). "The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core." In: *Quarterly Journal of the Royal Meteorological Society* 141.687, pp. 563–579. DOI: 10.1002/qj.2378.

EIDESSTATTLICHE VERSICHERUNG
DECLARATION ON OATH

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

I hereby declare upon oath that I have written the present dissertation independently and have not used further resources and aids than those stated.

Hamburg, den 25. November 2022
Hamburg, 25 November 2022

Theresa Claudia Lang