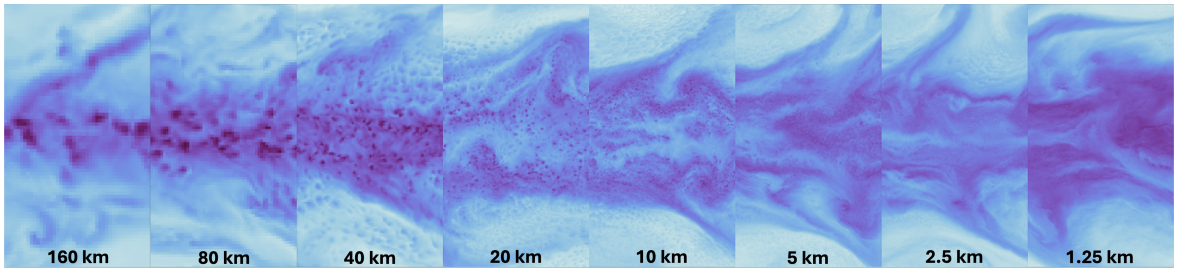
MAX-PLANCK-INSTITUT
FÜR METEOROLOGIE

Convergence of the Simulated Tropical Convergence Zone and its Response to Climate Change with Resolution



Angel Peinado Bravo

Hamburg 2025

Berichte zur Erdsystemforschung
Reports on Earth System Science

291
2025

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.

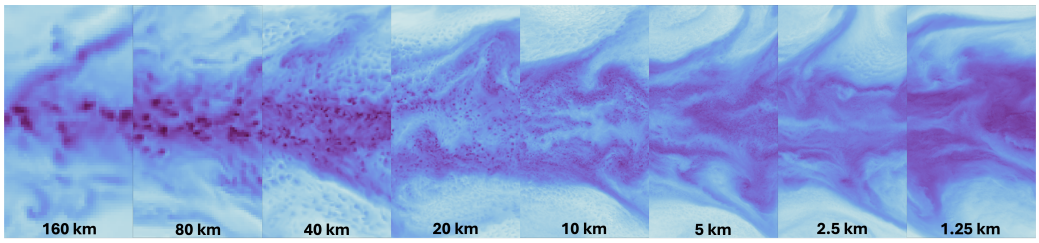
The "Reports on Earth System Science" continue the former "Reports" and "Examensarbeiten" of the Max Planck Institute.

Layout

Bettina Diallo and Norbert P. Noreiks
Communication

Copyright

Convergence of the Simulated Tropical Convergence Zone and its Response to Climate Change with Resolution



Angel Peinado Bravo

Hamburg 2025

Angel Peinado Bravo
aus Lima, Peru

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

Tag der Disputation: 9. April 2025

Folgende Gutachter empfehlen die Annahme der Dissertation:
Prof. Dr. Bjorn Stevens
Dr. Daniel Klocke

Vorsitzender des Promotionsausschusses:
Prof. Dr. Hermann Held

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

Titelgrafik: *Instantaneous precipitable water in the intertropical convergence zone (30N30S) across aquaplanet experiments with different horizontal grid spacing. The horizontal grid spacing is progressively halved from 160 to 1.25 km.*
Credit: Angel Peinado Bravo

ISSN 1614-1199 - doi: 10.17617/2.3661744

To my parents, of whose authentic gray hair I am one of the main causes.

ACKNOWLEDGMENTS

This dissertation marks the end of a voyage shaped by laughs, a renewed sense of discovery, and unexpected events. Along the hills and valleys, many helped me take one step after another, and I will forever cherish their complicity.

First of all, words cannot express my gratitude to my supervisor, Daniel Klocke. His positive way of viewing the impossible possible helped me overcome every cliff I encountered, even if I needed to jump into the unknown of ICON code and crush many supercomputers. Working with him made every step more enjoyable and full of new insights. If we did not find a way around a ridge, Bjorn Stevens was there to show us the way ahead, and I owe many thanks to him. His vision to flip a problem into an adventure without losing focus on the main question inspired me to challenge myself even more. I would also like to thank my panel chair, Peter Korn, for giving me encouraging words. Every panel meeting was enjoyable rather than challenging, thanks to him.

I am grateful to Jiawei Bao, Hans Segura, and Sebastian Ortega, who helped me test my ideas through laughs and long talks. I would also like to thank Cathy Hohennege, who welcomed me to the Climate Surface Interaction (CSI) group and provided invaluable feedback along this voyage. I appreciate Rene Reddler's and Jan Frederik Engels's countless discussions about ICON and their guidance in me through Levante's early, untamed days—the "Levante wild-west times." I am eternally grateful to Angela Gruber, Jacqueline Behncke, Danai Filippou, Maria-Jesus Rapanague, Jairo Segura, Sandro Rivas, and Daniela Ayvazova for their emotional support and amusement they brought to my life. In addition, I need to thank Lukas Kluft, Monika Esch, Dian Putrasahan, Hauke Schmidt, Arim Yoon, Clara Bayley, Hernan Campos, and everyone else from the CPH and the institute. Antje Weitz, Florian Mundt, and the IMPRS office also deserve special thanks for always being there and helping me through any problem I encounter with bureaucracy.

Finally, this would not have been possible without the support of my parents and family, especially my dog, Pimienta, as they are the reason I was able to start this journey in the first place.

ABSTRACT

More than 50 years have passed since the first attempts to understand the atmospheric general circulation in a comprehensive climate model. General circulation models (GCMs) have played an essential role in modeling and understanding the atmospheric general circulation and its response to climate change. However, their construction (horizontal grid resolutions in the order of 100 km and subgrid processes representation through parameterization) has raised criticism as uncertainties in global warming have remained almost the same along the different Coupled Model Intercomparison Projects (CMIPs, e.g., Meehl et al., 2020; Zelinka et al., 2020). One of the main contributors to inter-model spread in global warming in response to increased CO_2 concentration is the change of clouds, attributable to differences in cumulus parameterizations. However, clouds are complex features of the atmosphere that span a multi-scale range of processes, which makes it difficult to adequately represent physically at horizontal grid resolutions in the order of 100 km. To address the underlying uncertainty, more complex parameterization or explicitly resolving convection at adequate horizontal grid spacing in the order of one kilometer can be used; the so-called convection-permitting models or global storm resolving models (GSRM). However, the computational cost to perform such global simulations increases. In this dissertation, I first explore, in a state-of-the-art GSRM, if, by increasing the horizontal resolution, the atmospheric general circulation displays physical convergence as uncertainties and discretization errors reduce. Secondly, I explore if its response to forcing, mimicking climate warming, and the responses to climate change of the atmospheric general circulation display convergence.

In the first part of this dissertation, I focus on evaluating physical convergence of the atmospheric general circulation and increasing horizontal resolution using the ICOsahedral Nonhydrostatic (ICON) model from 160 km to 1.25 km horizontal grid spacing. I develop a methodology based on the Richardson extrapolation method to assess physical convergence in an idealized setup, which retain basic atmospheric features of Earth's general circulation, response to warming, and required reduced computational resources to achieve robust statistics. As I increase the horizontal grid spacing, a better representation of clouds and zonal distribution of water vapor drives convergence in the energy and water budget towards kilometer-scale [O(1 km)] horizontal grid spacing. The atmospheric general circulation displays convergence in its structure at kilometer-scale horizontal grid spacing, but its intensity requires finer horizontal grid spacing. However, shallow marine boundary layer clouds and their effect on the shortwave radiation components of the energy budget would require hectometer-scale horizontal grid spacing to achieve convergence, according to Large-Eddy Simulation (LES) findings (e.g., Stevens et al., 2020). ICON displays physical convergence with increasing horizontal resolution, and at 2.5-5 km horizontal grid spacing, uncertainties from numerical errors are significantly smaller in the large-scale structure of the general circulation.

In the second part of this dissertation, I investigate the model's response to a uniform increase in sea surface temperature, mimicking global warming. Across resolution and according to previous studies in GCMs, the hydrological cycle intensifies slower than the increase of precipitable water, the Hadley cell expands, and the anvil clouds shift to higher altitudes. The hydrological cycle intensification and longwave climate feedback

converge faster than other metrics at 10 km horizontal grid resolution, in agreement with observations and other GSRMs. The response of the width of the deep tropics varies according to the Intertropical Convergence Zone (ITCZ) structure, a contraction for a single ITCZ, and an expansion otherwise, converging at 5 km horizontal grid resolution. Across resolution, the boundary layer becomes drier while the middle and upper troposphere become moister and significantly warmer. The amount of anvil clouds in the deep tropics and shallow clouds in the subtropics do not show significant changes. My results show that ICON displays robust responses to warming with increasing horizontal resolution and converging for most large-scale features at 5 km, even if shallow marine boundary layer clouds have not converged at those resolutions.

In summary, this dissertation demonstrates that a GSRM (ICON) displays convergence with increasing horizontal grid spacing. Thus, the results raise confidence in using GSMR to investigate the climate response to warming. However, if one GSRM converges, it does not imply that another will or that it arrives at the same climate. For this reason, the methodology developed serves as a tool to assess GSRM implementations and further development.

ZUSAMMENFASSUNG

Seit den ersten Versuchen, die allgemeine atmosphärische Zirkulation in einem umfassenden Klimamodell zu verstehen, sind mehr als 50 Jahre vergangen. Allgemeine Zirkulationsmodelle (GCMs) haben eine wesentliche Rolle bei der Modellierung und dem Verständnis der allgemeinen atmosphärischen Zirkulation und ihrer Reaktion auf den Klimawandel gespielt. Ihre Konstruktion (horizontale Gitterauflösung in der Größenordnung von 100 km und Darstellung von Prozessen in Teilgittern durch Parametrisierung) hat jedoch Kritik hervorgerufen, da die Unsicherheiten in Bezug auf die globale Erwärmung bei den verschiedenen Coupled Model Intercomparison Projects (CMIPs, e.g., Meehl u. a., 2020; Zelinka u. a., 2020) fast gleich geblieben sind. Eine der Hauptursachen für die Streuung zwischen den Modellen bei der globalen Erwärmung als Reaktion auf eine erhöhte CO_2-Konzentration ist die Veränderung der Wolken, die auf Unterschiede bei der Parametrisierung der Kumuluswolken zurückzuführen ist. Wolken sind jedoch komplexe Merkmale der Atmosphäre, die eine Vielzahl von Prozessen umfassen, was es schwierig macht, sie bei horizontalen Gitterauflösungen in der Größenordnung von 100 km adäquat darzustellen. Um die zugrundeliegenden Unsicherheiten zu berücksichtigen, können komplexere Parametrisierungen oder die explizite Auflösung der Konvektion mit einem angemessenen horizontalen Gitterabstand in der Größenordnung von einem Kilometer verwendet werden; die so genannten konvektionszulassenden Modelle oder globalen sturmauflösenden Modelle (GSRM). Der Rechenaufwand für die Durchführung solcher globalen Simulationen steigt jedoch. In dieser Dissertation untersuche ich zunächst in einem hochmodernen GSRM, ob die allgemeine atmosphärische Zirkulation durch Erhöhung der horizontalen Auflösung physikalische Konvergenz zeigt, wenn Unsicherheiten und Diskretisierungsfehler abnehmen. Zweitens untersuche ich, ob die Reaktion der allgemeinen atmosphärischen Zirkulation auf die Erwärmung des Klimas und die Reaktionen auf den Klimawandel Konvergenz zeigen.

m ersten Teil dieser Dissertation konzentriere ich mich auf die Bewertung der physikalischen Konvergenz der allgemeinen atmosphärischen Zirkulation und die Erhöhung der horizontalen Auflösung unter Verwendung des ICOsahedral Nonhydrostatic (ICON)

Modells von 160 km auf 1,25 km horizontalen Gitterabstand. Ich entwickle eine Methodik auf der Grundlage der Richardson-Extrapolationsmethode, um die physikalische Konvergenz in einem idealisierten Aufbau zu bewerten, der die grundlegenden atmosphärischen Merkmale der allgemeinen Erdzirkulation und die Reaktion auf die Erwärmung beibehält und weniger Rechenressourcen benötigt, um robuste Statistiken zu erhalten. Wenn ich den horizontalen Gitterabstand vergrößere, führt eine bessere Darstellung der Wolken und der zonalen Verteilung des Wasserdampfs zu einer Konvergenz des Energie- und Wasserhaushalts in Richtung eines horizontalen Gitterabstands auf Kilometerebene [O(1 km)]. Die allgemeine atmosphärische Zirkulation zeigt Konvergenz in ihrer Struktur bei einem horizontalen Gitterabstand auf Kilometerskala, aber ihre Intensität erfordert einen feineren horizontalen Gitterabstand. Die flachen marinen Grenzschichtwolken und ihre Auswirkungen auf die kurzwelligen Strahlungskomponenten des Energiehaushalts würden jedoch eine horizontale Gitteraufteilung auf Hektometerskala erfordern, um Konvergenz zu erreichen, so die Ergebnisse der Large-Eddy-Simulation (LES, e.g., Stevens u. a., 2020). ICON zeigt physikalische Konvergenz mit zunehmender horizontaler Auflösung, und bei einem horizontalen Gitterabstand von 2,5 bis 5 km sind die Unsicherheiten durch numerische Fehler in der großräumigen Struktur der allgemeinen Zirkulation deutlich geringer.

Im zweiten Teil dieser Dissertation untersuche ich die Reaktion des Modells auf einen einheitlichen Anstieg der Meeresoberflächentemperatur, der die globale Erwärmung nachahmt. Bei allen Auflösungen und entsprechend früherer Studien in GCMs intensiviert sich der Wasserkreislauf langsamer als der Anstieg des niederschlagbaren Wassers, die Hadley-Zelle dehnt sich aus, und die Ambosswolken verlagern sich in höhere Lagen. Die Intensivierung des hydrologischen Zyklus und die langwellige Klima-Rückkopplung konvergieren bei einer horizontalen Gitterauflösung von 10 km schneller als andere Messgrößen, was mit Beobachtungen und anderen GSRMs übereinstimmt. Die Reaktion der Breite der tiefen Tropen variiert je nach der Struktur der Intertropical Convergence Zone (ITCZ), eine Kontraktion bei einer einzigen ITCZ und ansonsten eine Ausdehnung, die bei einer horizontalen Gitterauflösung von 5 km konvergiert. Mit zunehmender Auflösung wird die Grenzschicht trockener, während die mittlere und obere Troposphäre feuchter und deutlich wärmer wird. Die Menge der Ambosswolken in den tiefen Tropen und der flachen Wolken in den Subtropen zeigen keine signifikanten Veränderungen. Meine Ergebnisse zeigen, dass ICON mit zunehmender horizontaler Auflösung robuste Reaktionen auf die Erwärmung zeigt und für die meisten großräumigen Merkmale bei 5 km konvergiert, auch wenn die flachen Wolken der marinen Grenzschicht bei diesen Auflösungen nicht konvergieren.

Zusammenfassend lässt sich sagen, dass diese Dissertation zeigt, dass ein GSRM (ICON) mit zunehmendem horizontalen Gitterabstand Konvergenz zeigt. Die Ergebnisse erhöhen somit das Vertrauen in die Verwendung von GSMR zur Untersuchung der Reaktion des Klimas auf die Erwärmung. Wenn ein GSRM konvergiert, bedeutet dies jedoch nicht, dass ein anderes auch konvergiert oder dass es zum gleichen Klima kommt. Aus diesem Grund dient die entwickelte Methodik als Instrument zur Bewertung von GSRM-Implementierungen und Weiterentwicklungen.

ACRONYMS

AMIP . . . . . Atmospheric Model Intercomparison Project

CFD . . . . . Computational Fluid Dynamic

CFL . . . . . Courant–Friedrichs–Lewy

CMIP . . . . . Coupled Model Intercomparison Project

COOKIE . . . . . Clouds On-Off Klimate Intercomparison Experiment

DYAMOND . . . . . DYnamics of the Atmospheric general circulation Modelled On Non-hydrostatic Domains

ECS . . . . . Equilibrium Climate Sensitivity

FLOPS . . . . . Floating-Point operations per second

GCM . . . . . General Circulation Model

GSRM . . . . . Global Storm Resolving Model

ICON . . . . . ICOsahedral Nonhydrostatic model

ITCZ . . . . . Intertropical Convergence Zone

NICAM . . . . . INon-hydrostatic ICosahedral Atmospheric Model

PDE . . . . . Partial Differential Equation

RCE . . . . . Radiative-Convective Equilbrium

RCEMIP . . . . . Radiative-Convective Equilbrium Model Intercomparison Project

SST . . . . . Sea Surface Temperature

CONTENTS

Part I

UNIFYING ESSAY

1

INTRODUCTION AND MOTIVATION

The walls of this chamber are painted to form a map of the globe. The ceiling represents the north polar regions.... A myriad of computers are at work upon the weather of the part of the map where each sits.

- Lewis Fry Richardson (Richardson, 2007)

Back in the winter of 1917, Richardson Fry Lewis began a journey into weather forecasting, armed with pencil, paper, a slide rule, and a table of logarithms. Richardson sought to forecast central European weather over a span of six hours, using sparse observational data (Richardson, 2007). The endeavor failed and fell short due to space and time numerical discretization, and initial condition errors (Lynch, 1992); however, he published his results and ended his book with a vision of an orchestra - a hall of human computers, echoing with the click of slide rules and the quiet murmur of humans thoughts weaving weather predictions. Richardson set the gears of numerical models for weather prediction at the right time as the first programmable general-purpose computers began to emerge in the 1940s.

At the Institute of Advances Studies in Princeton, Jule Charney and John von Neumann ventured upon a vibrant revolution. They used ENIAC, the first digital computer with a performance of just 500 Floating Point Operations per Second (FLOPS; just $2 \cdot 10^9$ times slower than an iPhone XS), to solve the barotropic vorticity equation (J. G. Charney and Neumann, 1950), starting the revolution of climate computing. In 1955, at the same university, Norman Phillips challenged the power of computers to simulate the gross features of the atmospheric general circulation (Phillips, 1956); even though he simplified the globe as a cylinder, he showed how large eddies played a key role in atmospheric energy and momentum transport, creating the first 'true' General Circulation Model (GCM). In the following years, Smagorinsky invited Manabe to join his laboratory and changed the history of GCMs and climate science. Together, they incorporated more physics into their model, i.e., the radiation effect by water vapor, ozone, and CO_2, creating the first three-dimensional model that solves basic equations for a global simplified atmosphere (Manabe and Bryan, 1969). In this way, they developed the first model that coupled the atmospheric and ocean response to anthropogenic CO_2 emissions and highlighted the unequivocal result of climate change (Manabe and Bryan, 1969). Since then, climate models have grown in complexity backed by the exponential growth of available computing power. Many research centers incorporated more sophisticated Earth system components into their models through sub-grid scale processes, ready to devour every clock cycle and byte of memory.

Climate science and models have a long history, with over 100 years of progress in understanding the atmosphere's general circulation. Undoubtedly, GCMs have been an important tool for understanding climate change (i.e., Held and Soden, 2006; Voigt and Shaw, 2015) with a typical horizontal grid spacing of $\sim$100 km. However, GCMs have long-standing unresolved biases, i.e., double ITCZ bias, and no model-tuning approach

or complex parameterization has been able to resolve them. Therefore, it might not lead to unraveling the emergent question of climate change (Palmer and Stevens, 2019), i.e., regional response. Thus, we need a new strategy to solve the emergent questions for climate change by resolving key features at an appropriate horizontal grid spacing [O(1km)], i.e., deep convection and ocean mesoscale eddies. To bridge this gap, Global Storm Resolving Models (GSRMs) emerged with the era of supercomputers, and the development of the Non-hydrostatic ICosahedral Atmospheric Model (NICAM; Tomita et al., 2005; Miura et al., 2007b; Sato et al., 2008) in Japan and, in parallel, the ICOsahedral Nonhydrostatic (ICON; Gassmann and Herzog, 2008; Wan et al., 2013; Zängl et al., 2015) in Germany. GSRMs with kilometer-scale [O(1km)] horizontal grid spacing and less parameterizations resolve convective processes and the associated vertical energy transport, which is the dominant mode of energy transport in the tropics, representing the multi-scale nature of convection. GSRMs have shown many improvements compared to GCMs, i.e., tropical cyclones (i.e., Judt et al., 2021; Baker et al., 2024), the Madden-Julian Oscillation (i.e., Miura et al., 2007a; Miyakawa et al., 2014; Takasuka and Satoh, 2021), precipitation diurnal cycle (i.e., Ma et al., 2022; Song et al., 2024), and mesoscale convective systems (i.e., Feng et al., 2023) representation. Nevertheless, the increase in horizontal grid spacing does not solve all the problems and faces many challenges.

The most evident challenge of GSRMs is the increased computational cost to achieve global kilometer or sub-kilometer scale simulations. As we increase the horizontal grid spacing, we must reduce the time step to stabilize the model (CFL restriction), thus making the simulation expensive by a factor of eight if we refine the horizontal grid spacing by a factor of two. The appearance of Exa-scale supercomputers (> 1e18 FLOPS) in the last years brought the computational capabilities to simulate 1 year per day at 1.25 km horizontal grid spacing (Giorgetta et al., 2022), even though it would require Exa-scale supercomputer total capacity to run it. Even if we can run such a model, an underlying question arises: How does the increase of horizontal grid spacing with resolved meso-scale affect the large-scale structures or the climate state? Do we observe physical convergence with increasing horizontal grid spacing?

Before drafting the research gap on the physical convergence of GSRM and its response to global warming, we need to dip into some concepts in the following section.

1.1 BACKGROUND

This section reviews how convergence has been understood in climate models, in what type of simulations convergence has been investigated, and lastly, a brief review of remaining models uncertainties in climate response to warming.

1.1.1 *Physical convergence in climate models*

From mathematics and physics perspective, a climate system can be represented and summarized into a set of non-linear partial differential equations (PDE) of the general form:

$$\frac{\partial}{\partial t}\mathbf{X} + \mathbf{V} \cdot \nabla \mathbf{X} + \mathbf{D}\left(\mathbf{X}\right) = \mathbf{S}\left(\mathbf{X}, t\right) \tag{1.1}$$

where, $\mathbf{X}$ is the vector of atmospheric prognostic variables, $\mathbf{V}$ is the velocity vector, $\mathbf{D}$ are the interactions within $\mathbf{X}$ variables, and $\mathbf{S}$ is the source or sink term (sub-grid scale physical processes, i.e., heating due to mixing in the thermodynamic equation). To solve this set of equations and develop a climate model, we need to make some decisions, i.e., how to discretize the space and the numerical method to solve the partial differential equations (i.e. Staniforth and Thuburn, 2012). According to our decision, we introduce numerical errors in our model, and we need to verify if the model has three main properties: convergence, consistency, and stability. A model is stable if the truncation error decays with each time step; is convergent if, by reducing the spatial discretization, it approaches asymptotically to some fixed value (numerical error reduction); and is consistent if, by reducing the spatial and time discretization, the truncation reduces and the solution approaches the "true" solution of the PDE. There is no easy way to verify these properties in complex non-linear partial differential equations unless we run the model.

To verify convergence and the effect of increasing resolution in a climate model, let's first use the operator $()_{\mathrm{R}}$ to equation 1 as performed by Boer and Denis (1997):

$$squm\frac{\partial \mathbf{X}_{\mathrm{R}}}{\partial t} + \mathbf{V}_{\mathrm{R}} \cdot \nabla \mathbf{X}_{\mathrm{R}} = \mathbf{S}\left(\mathbf{X}_{\mathrm{R}}, t\right) + \mathbf{S}_{\mathrm{D}}\left(\mathbf{X}_{\mathrm{R}}, t\right) + \mathrm{E}_{\mathrm{S}} + \mathrm{E}_{N} \tag{1.2}$$

where, $\mathbf{X}_{\mathrm{R}}$ are the resolved components of $\mathbf{X}$, $\mathbf{S}\left(\mathbf{X}_{\mathrm{R}}, t\right)$ are the sub-grid scale processes (parameterizations) of the resolved scales, $\mathbf{S}_{\mathrm{D}}\left(\mathbf{X}_{\mathrm{R}}, t\right)$ are the parameterization of the interaction between resolved and unresolved scales, E_{S} is the parameterization error, and E_{N} is the numerical error from time and space discretization. In climate models, the usage of parameterizations obscures convergence analysis since numerical and parameterization errors cannot be separated because they depend on resolution. Thus, an increase in resolution might lead to non-convergent behavior, deterioration of simulated climate, or only minimal improvement (Williamson, 2008).

But then, how do climate models have been verified? Most climate models in the realm of GCMs have avoided the question of convergence. Climate models have been compared to observations or within each other in intermodel comparison projects, such as the Coupled Model Intercomparisson Projects (CMIPs, i.e. Chen and Frauenfeld, 2014; Koutroulis et al., 2016). Now, it is expected to define that a "good" climate model resembles some climate characteristics, i.e., the increase in sea surface temperature in the last decade, and lies within an intermodel spread. However, how well a model resembles the current climate does not translate to the climate change response (Klocke et al., 2011; Parker, 2018), and it raises the question of the sources of intermodel differences in CMIPs (Zelinka et al., 2020). Zelinka et al. (2022) compared the response and feedback to global warming of cloud processes (cloud feedback) by expert assessments (Sherwood et al., 2020) and found that models display biases from diverse cloud feedback components from parameterization, i.e., cumulus parameterization, rather than a unique source. Moreover, models that scored well according to expert assessment do not guarantee a more skillful simulation of cloud feedback. In contrast, GSRMs use less complex parameterization at the expense of increasing resolution; thus, reducing the interplay between numerical and parameterization error is expected. Additionally, increasing the horizontal grid resolution improves external fields of the climate model, particularly orography (i.e., Prein et al., 2016), which influences climate processes, i.e., precipitation (i.e., Oouchi et al., 2009; Langhans et al., 2012).

In the Computational Fluid Dynamics (CFD) community, model results are accepted by showing the reduction of truncation errors and increased accuracy by systematically increasing resolution (Roache et al., 1986). Even though there is no exact protocol, J. Freitas (1993) proposed guidelines for reporting uncertainties due to discretization in the CFD community, which is known as a grid-convergence study. However, in climate models, these guidelines have not been used and avoided the question of convergence directly (Williamson, 2008). Instead, a sensitivity study of resolution or how refinement affects the climate has been preferred. A common practice is to compare a quantity or characteristic of the finest to the coarser resolution (i.e. Williamson, 2008; Landu et al., 2014; Hohenegger et al., 2020); this method assumes that the finest resolution is closer to the truth as the numerical error reduces as we increase resolution. However, a drawback of this approach lies in how confident can we be that the finest resolution is actually converging or closer to the truth. Regardless, this approach can provide a hint of the convergence regime and information, if and only if the quantity converges within the resolutions studied. Interestingly, for convergence studies, some papers compare quantitative differences through global means and selected metrics and other qualitatively regional responses or structures, i.e., zonal structures and snapshots (i.e., Williamson, 2008; Langhans et al., 2012; Landu et al., 2014; Zarzycki et al., 2019). However, it leaves much room for the reader to interpret convergence. Hohenegger et al. (2020) used another approach in addition to the previous one by comparing resolution differences to the DYAMOND inter-comparison project intermodel spread (Stevens et al., 2019) for different quantities. Even though the DYAMOND simulations are too short to observe climate states climatology, and results might lie within spin-up and inter-variability, they serve to understand more transient climate features, i.e., cyclones (Judt et al., 2021; Pantillon et al., 2024), and the nature of cloud morphology (Freischem et al., 2024). Thus, is there a more quantitative way to evaluate convergence in climate models by increasing resolution? This gap is addressed in section 2.2, where we propose and develop an alternative to evaluate convergence.

1.1.2 *Convergence and simulation hierarchies*

In the pursuit of understanding our climate, models have become worlds within worlds, evolving into one of the most complex computer programs nowadays. Held (2005) highlighted the need to develop a series of hierarchies of climate simulations to close the gap between our theoretical conceptualization of climate processes and comprehensive climate models. Yet, there is a wide range of experiments across the development of climate models. Echoing Held's (2005) vision, if convergence with increasing resolution is considered a property of climate models, a hierarchy of experiment types is required to answer which experiment type might be ideal to understand convergence.

Held (2005) stressed the need to keep models within the hierarchy as *elegant*; in other words, they should have the lowest complexity for the question to be answered. An outcome of the "Model Hierarchy Workshop 2016" was Figure 1.1 (Jeevanjee et al., 2017), which tries to encase *elegant* models (blue color) in a 2D diagram over three-axis: bulk (convection and radiation), boundary (ocean and land), and dynamics (rotation and fluid) forcings. However, *elegant* tends to be subjective, and Jeevanjee et al. (2017) wondered whether resolving convection without parameterization is potentially *elegant*.

By design, GSRMs become *elegant* by reducing the degrees of freedom and complexity within parameterizations and resolving convection in their multi-scale nature, reducing the interplay between parameterization and numerical errors. For instance, Zarzycki

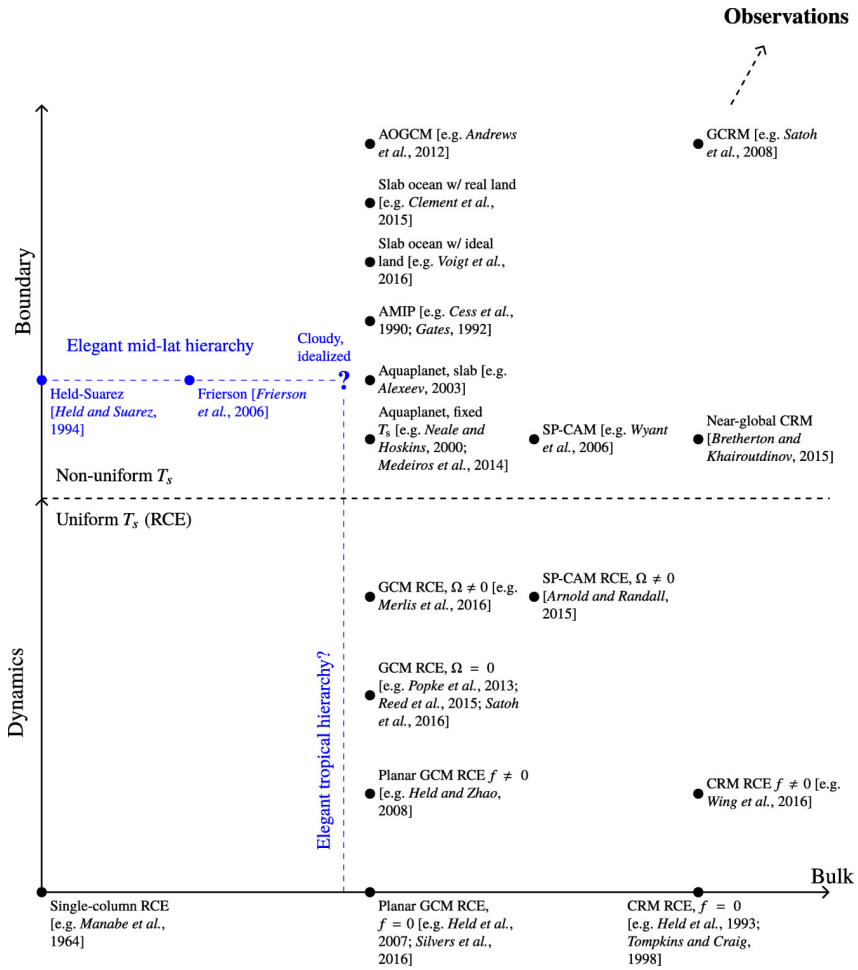


Figure 1.1: Climate simulation hierarchy, x-axis, bulk forcing (mainly radiation), y-axis, split into two pieces, dynamic (mainly rotation) and boundary (mainly SST, uniform or non-uniform). Ω is the angular velocity, and f the Coriolis parameter. Blue highlights the *elegant* midlatitude models (Held-Suarez and Frierson) and the boundary of possible *elegant* models in the tropics. The blue lines terminate in a nonexistant *elegant* model that incorporates midlatitudes, tropics, and interactive clouds. The illustration is taken from Jeevanjee et al. (2017)

et al. (2019), through the Dynamical Core Model Intercomparison Project (DCMIP2016), investigated the horizontal grid spacing effect (horizontal resolution convergence) of the splitting supercell test case (long-lived convective cell). They observed a sensitivity reduction across models, including GSRMs, as they approached 500 m horizontal grid spacing of bulk properties. Meanwhile, precipitation patterns of the last time step (snapshot) displayed a weaker difference reduction with increasing resolution. Strikingly, the inter-model spread was more significant even at 500 m horizontal resolutions, suggesting that differences within the model's physics implementation partly determine the convergent solution of the transient dynamics. The splitting supercell test case does not answer how these small differences might be relevant for climate. Increasing the complexity of the experiment, Langhans et al. (2012) explored the statistical convergence in regional climate models in the European alpine region from 4.4 to 0.55 km horizontal grid spacing for a period of nine simulated days. They observed a convergence of bulk properties and the diurnal cycle of precipitation with increasing horizontal grid spacing. Additionally, Vergara-Temprado et al. (2020) investigated at which resolutions convection

is explicitly resolved using a regional climate model. They found that deactivating the cumulus parameterization has a better diurnal precipitation cycle representation for finer horizontal grid spacing than 50 km, comparable to increasing the horizontal grid resolution. Nevertheless, the regional climate model convergence is influenced by its boundary conditions, and the small-scale to large-scale interaction is missing.

In the framework of global scale simulations, Kajikawa et al. (2016) explored the dependency of deep convection to resolution using NICAM with realistic boundary conditions and horizontal grid spacing going from 14 to 0.86 km for 12 simulated hours, similar to the protocol of the Atmospheric Model Intercomparison Project (AMIP). The study revealed an apparent convergence at about 1 km horizontal grid spacing for convective cell numbers and no resolution dependency for global statistics. Using a similar setup, Hohenegger et al. (2020) used ICON in the DYAMOND inter-comparison project configuration (Stevens et al., 2019) with horizontal grid spacings varied from 80 to 2.5 km to investigate physical convergence over 40 simulated days. They quantified convergence by comparing horizontal resolution differences to the DYAMOND inter-model ensemble spread, observing a convergence towards 2 km horizontal grid spacing for bulk and general circulation characteristics; however, the differences across models participating in the DYAMOND project are still substantial, reflecting that 40 days cannot capture climate statistics robustly, or that the models converge to statistically different climates. In a separate study, Schmidt et al. (2023) explored the effects of vertical resolution and showed that there was surprisingly little sensitivity to the vertical grid spacing for grid spacing finer than the standard configuration of 90 stretched vertical levels with model top at 75 km.

Even if AMIP simulations are desired to investigate the atmosphere component of climate models, they require a long integration time of about 30 years, to evaluate climate statistics. In the framework of kilometer-scale [O(1km)] simulation, 30 years will need a large amount of computational resources. We must find a compromise to minimize computational cost for convergence studies. We can reduce the experiment's complexity by using aquaplanet experiments. Aquaplanet experiments preserve basic features of Earth's general circulation (Blackburn et al., 2013; Medeiros et al., 2015, 2016) while offering a simplified framework without complex boundary interactions and they have proven themselves as useful tools to provide valuable insights into phenomena such as tropical circulation and variability (Möbis and Stevens, 2012; Medeiros et al., 2016; Popp et al., 2020; Rios-Berrios et al., 2020), how clouds respond to warming (Voigt and Shaw, 2015; Talib et al., 2018), and the response of other aspects of the climate to external forcing (Stevens and Bony, 2013a; Medeiros et al., 2015; Retsch et al., 2019). Medeiros et al. (2015) highlighted the usage of aquaplanet experiments by comparing their response to warming to AMIP simulations, showing the response of main characteristics, i.e., climate feedback and circulation response, are similar or equal. Thus, they pose great candidates for understanding convergence in GSRMs with and without forcing in a simplified framework, which requires a reduced simulation time to achieve robust statistics. I will further detail the aquaplanet experiments characteristics in section 2.1.

1.1.3 *Uncertanities in Climate Change*

More than 40 years ago, Charney published a landmark report on "Carbon Dioxide and Climate: A Scientific Assessment," known as the Charney Report (Charney, 1979). Although they did not have access to advanced computer facilities, multi-decadal satellite observations of global climate change, and numerous sophisticated climate models, their

main results are still relevant and have aged well. Their conclusion regarding the long-term rise in global mean temperature (near-surface air) expected from a doubling of atmospheric CO_2: "We estimate the most probable global warming for a doubling of CO_2 to be near 3°C with a probable error of 1.5°C," is remarkably close to the last CMIP estimates (i.e. Meehl et al., 2020; Zelinka et al., 2020, 2022). The temperature increase associated with an increase of CO_2, the Equilibrium Climate Sensitivity (ECS), is estimates as changes in global mean energy balance at the top of the atmosphere (TOA) reaches the equilibrium state following a doubling of CO_2. At equilibrium, $ECS = -F/\lambda$, where changes in global mean energy balance at TOA are decomposed into the radiative forcing (F) and the climate feedback parameter (λ). Meehl et al. (2020) reported that across the different CMIPs, the ECS range has almost stayed the same, and in the last CMIP6, the ECS range has been the largest across the 26 evaluated models. Another relevant conclusion in Charney's report is the critical role of cloud processes and feedback in estimating ECS. Zelinka et al. (2020) investigated the main causes of uncertainties within CMIP6 models, showing that the high climate sensitivity within models is attributed to differences in the cloud response to warming. In particular, large uncertainties were observed in the reduction of the albedo effect of clouds in middle and high latitudes, as well as in changes to tropical high clouds. Clouds are not only simple trackers of large-scale circulation, but their radiative effect also plays a mayor role in setting the large-scale circulation (Bony et al., 2015). Their interplay with the climate response is complex and has raised different intercomparison projects. The Clouds On Off Klima Intercomparison Experiment (COOKIE, Stevens et al., 2012) focuses on the cloud radiative heating and cooling in the large-scale circulation by making clouds transparent to the radiation component (Voigt et al., 2021). One of the most interesting results is the relevance of the cloud-radiative effect in the Hadley cell expansion with warming. Without changes in the cloud-radiative effect, there would be hardly any poleward migration of the Hadley cell edge, signaling the higher relevance of cloud feedback in a regional rather than the global response (Voigt et al., 2021). The large uncertainty in cloud responses arises from differences in the representation of convection by cumulus parameterizations, as it is difficult to model cloud processes sufficiently at GCMs horizontal grid resolutions. As Stevens and Bony (2013a) highlighted, there are many inter-model differences in the cloud response and its effect on large-scale circulation, even in simplified aquaplanet experiments (Figure 1.2).

With a better physical representation of convection and clouds, GSRMs offer an opportunity to reduce the uncertainties and resolve the effects of small-scale and large-scale interaction (Voigt et al., 2021). Kodama et al. (2015) performed one of the first global AMIP simulations without convection parameterizations using NICAM at 14 km horizontal grid spacing with an integration time of 20 years. Using their simulation, Noda et al. (2019) found that the shortwave climate feedback displayed a neutral response; thus, the atmospheric global mean shortwave radiation do not show significant differences to warming. More recently, Merlis et al. (2024b) reported similar results using X-SHIELD, with a horizontal grid spacing of 3 km and an integration time of 2 years. Interestingly, CMIP models display large uncertainities in the shortwave climate feedback, as changes in shallow cloud amounts with a cooling effect (high albedo) varies between models. However, this might pose the question of whether the muted effect of shortwave climate feedback is affected by horizontal grid spacing as shallow clouds reduce with increasing resolution (Kajikawa et al., 2016; Hohenegger et al., 2020). Another relevant cloud response is the increase of anvil clouds in the tropics, the Iris effect (Lindzen et al., 2001; Bony et al., 2016). Silvers et al. (2023) used the Radiative-Convective Equilbrium Model

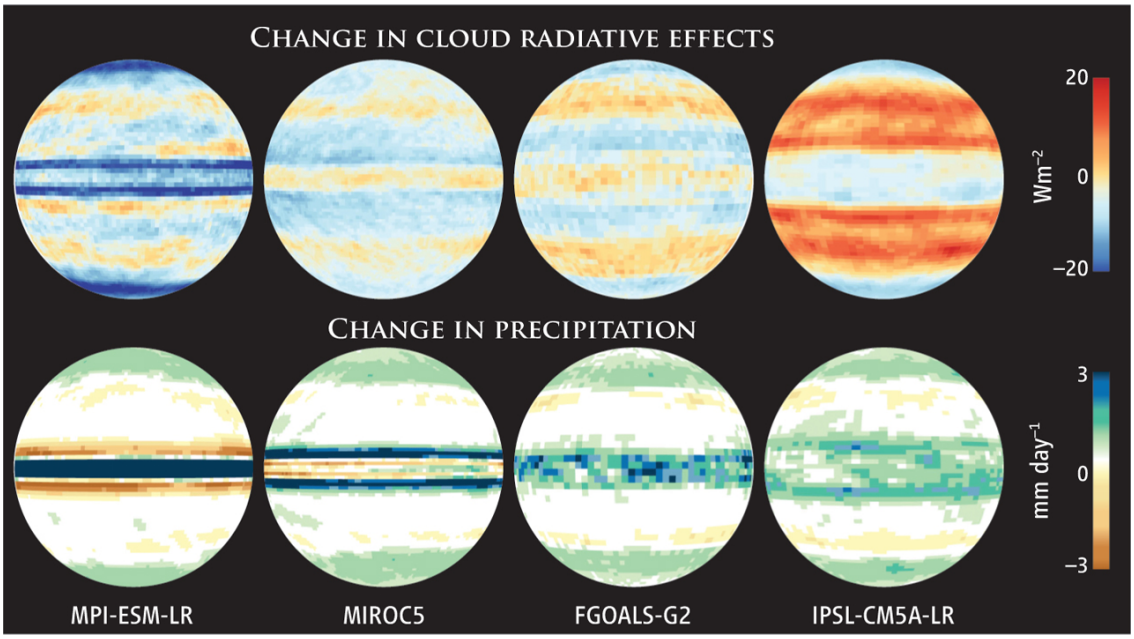


Figure 1.2: Response in aquaplanet experiments mimicking climate change by increasing the sea surface temperature by 4K using four different climate models (GCMs). The top row displays change in the atmospheric cloud radiative effect (ACRE) and bottom row, the change in precipitation. MPI-ESM-LR displays a contraction of the ITCZ as precipitation migrated equatorward, MIROC5 displays a poleward migration of the ITCZ, while FGOALS-G2 and IPSL-CM5A-LR don't show significant migration of the ITCZ. ACRE shows significant differences associated with circulation changes. The illustration is taken from Stevens and Bony (2013a)

Intercomparison Project (RCEMIP; Wing et al., 2018), which simulated the tropics in a simplified framework, and found that one-third of the models do not present the Iris effect. Neither Noda et al. (2019) nor Retsch et al. (2019) observed the Iris effect with their global simulations. Interestingly, Retsch et al. (2019) investigated the effect of resolution and convection (with and without cumulus parameterization) using aquaplanets and found that the climate feedback and ECS difference come from the different responses in the water vapor feedback instead of cloud responses. Nevertheless, it remains the question of how the response of the large-scale response to increased resolution, i.e., tropical circulation changes.

1.2 THE SCIENTIFIC PROBLEM: PHYSICAL CONVERGENCE IN A GSRM AND ITS RESPONSE TO GLOBAL WARMING

Having laid out the background and the current insights into convergence in climate models, I proceed to dip into the main research gaps that motivate this dissertation.

The first part of the dissertation focuses on developing a methodology to evaluate convergence qualitatively and quantitatively using aquaplanet experiments, which I detail in Chapter 2. I adopt the ICOsahedral Nonhydrostatic (ICON) model, which was initially developed in a partnership of the Max-Planck Institute of Meteorology (MPI-M) and the German Weather Service (DWD) and is now further developed and maintained in a larger consortium also including the German Climate Computing Centre (DKRZ), the Swiss Federal Institute of Technology (ETH), MeteoSwiss, and the Karlsruhe Institute for Technology (KIT), as the GSRM for this study. I use the atmosphere component of

ICON in a global aquaplanet experiment configuration, retaining only parameterizations for the radiant energy transfer, which is based on RTE-RRTMGP (Rapid Radiative Transfer model for General circulation model applications, Parallel, Pincus et al., 2019), a single moment bulk microphysical parameterization consisting of five condensate habits (Baldauf et al., 2011), and 3D turbulent mixing as described in Smagorinsky (1963) with modifications by Lilly (1962) as implemented in the ICON model by Lee et al. (2022a) following an earlier implementation by Dipankar et al. (2015). The vertical grid consists of 90 stretched levels, where levels are more finely spaced close to the surface than at the model top at 75 km, and a damping layer stretches from 44 km to the model top with damping increasing from 44 km upwards. In the following, I will refer to this configurations as ICON. I use this configuration to answer the following overarching questions:

RQ1. Does the GSRM ICON display convergence with increasing horizontal grid spacing?

To answer this question, I performed aquaplanet experiments with the same physics and vertical grid, varying only the horizontal grid spacing and the time step to keep the model stable. I use horizontal grid spacing from 160 to 1.25 km and focused on how resolving small-scale processes and their interaction with the large-scale circulation shows convergence with increasing horizontal grid spacing.

RQ2. If a GSRM shows convergence with increasing horizontal grid spacing, does its response to warming also show convergence?

To answer this question, I extended the simulation of the first study by 360, 180, and 90 days for experiments with horizontal grid spacing ranging from 160-80, 40-5, and 2.5 km, respectively, referred to as "control" in section 2. Using the same initial state, I performe forced simulations, mimicking climate warming by increasing the sea surface temperature (SST) by 4K uniformly for the same period, referred to as "forced" in section 2. For the "forced" experiments, I discard the first 45 days as the spin-up time. I disregarded the 1.25 km horizontal grid spacing experiment as it was not able to overcome the spin-up time due to computational constraints.

2

PHYSICAL CONVERGENCE METHODOLOGY

"The ability to simplify means to eliminate the unnecessary, so that the necessary may speak."

Hans Hofman

This chapter addresses how to determine convergence in a GSRM. Based on Chapter 1.1, we require a model or experiment type that captures atmospheric climate characteristics, retaining the interaction between tropics and extratropics circulations and clouds. For this purpose, we use aquaplanet simulations, which capture the main characteristics of the atmospheric general circulation. Additionally, its response to forcing is similar to AMIP simulations. Upon model selection, we require a methodology to assess the convergence of the main climate characteristics (global, large-scale, and regional). The methodology should incorporate estimates of convergence tendency and, ideally, a way to estimate which resolution might be required to achieve convergence. It should also incorporate uncertainty quantification for experiments that have yet to reach a statistically steady state.

In the following subsections, I further describe the aquaplanet's properties, the experiment setup, and the methodology developed to assess convergence.

2.1 AQUAPLANET PROPERTIES

Neale and Hoskins (2000) proposed a series of aquaplanet experiment setups as an intermediate step between experiments with complex boundaries (e.g., AMIP) and highly idealized experiments (e.g., single-column experiments), retaining the interaction within dynamics and parameterizations. Aquaplanet experiments model the ocean as a constant zonally symmetric sea surface temperature, the radiative forcing as a perpetual equinox (symmetric-constant irradiation about the equator) with a diurnal cycle, without sea-ice interaction, use well-mixed greenhouse gases, and a constant zonally symmetric ozone profile corresponding to the mean climatology used in AMIP simulations (Blackburn et al., 2013). Figure 2.1 displays the five different sea surface temperature profiles (SST-Profiles) proposed by Neale and Hoskins (2000). Rajedran et al. (2013) observed differences in the large-scale structure with different SST-profiles. About 90% of models that used SST-CONTROL display one single ITCZ, while SST-QOBS and SST-FLAT do not show agreement between models, displaying single or double ITCZ. Nevertheless, SST-QOBS experiments most closely resemble the atmospheric energy transport from observations (Williamson et al., 2013) and have been used frequently to investigate different climate characteristics, as described in section 1.1.2. Thus, I use SST-QOBS for our experiments, which follows the specification of latitudinal variations of:

$$T_s(\phi) = \begin{cases} \frac{27}{2}\left(2 - \sin^2\left(\frac{3\phi}{2}\right)\right)\left(1 + \sin^2\left(\frac{3\phi}{2}\right)\right) & \text{if} \quad |\phi| \leq \frac{\pi}{3} \\ 0 & \text{if} \quad \text{otherwise} \end{cases} \qquad [K] \qquad (2.1)$$

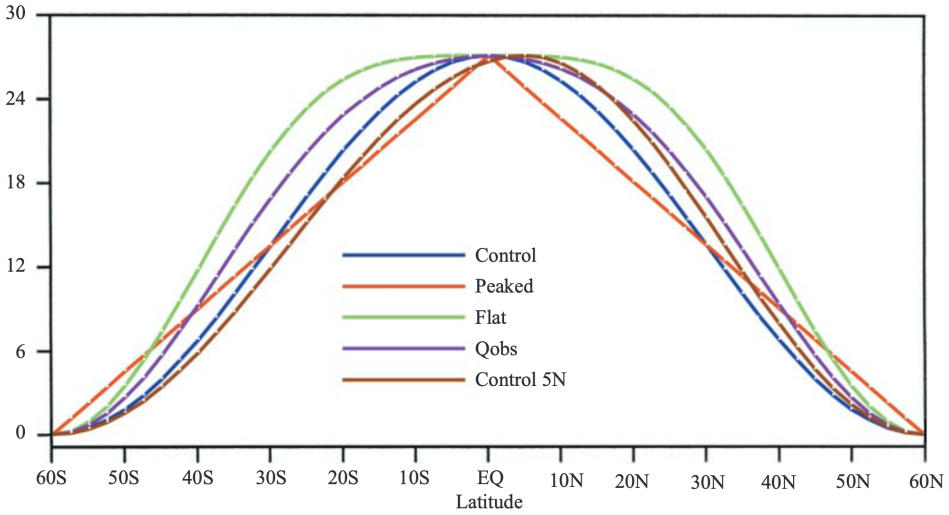


Figure 2.1: The five different sea surface temperature profiles (SST-Profiles) proposed by Neale and Hoskins (2000), SST-CONTROL, SST-PEAKED, QOBS, SST-FLAT, and SST-CONTROL-5N. The illustration is taken from Neale and Hoskins (2000)

By construction, an aquaplanet with either SST-QOBS, -CONTROL, -PEAKED, and -FLAT is symmetric with respect to the equator in its surface (SST-profile) and radiative forcing (perpetual equinox); thus, in its climate state (long integration time), the main climate characteristics, i.e., precipitation, should be symmetric to the equatorial axis or point symmetric to the equator, i.e., meridional wind. The symmetrical nature of aquaplanets is exemplified in Figure 2.2, which shows the mean precipitation over 180 days of integration time of an aquaplanet experiment at horizontal grid spacing of 5 km, its zonal mean, and the decomposition into its symmetrical and asymmetrical components. Since precipitation is symmetric to the equatorial axis, the symmetric component is the mean between the north and south hemispheres, while the asymmetrical component is the difference between the north and south hemispheres divided by two. The symmetric and asymmetric component definitions are reversed for variables that are point symmetric to the equator.

Neale and Hoskins (2000) noted that aquaplanets require integration times shorter than AMIP simulations as the symmetrical nature and the absence of complex boundary conditions require reduced sampling time to achieve the statistical steady state. More recently, Medeiros et al. (2016) investigated the sampling time for global and zonal structures with an estimated 24 months to achieve robust statistics. However, sampling times larger than 12 months are expensive for a global climate model with kilometer-scale horizontal grid spacing. I revisited the simulation time required to achieve robust statistics with a minimum simulation time by using the autocorrelation of successive daily means and define the sampling interval for the cumulative mean to approach a stationary value. We can be certain of the statistics once a daily or cumulative mean becomes decorrelated. However, autocorrelation is suitable for sampling statistics of scalar quantities, it does not capture vectors well, i.e., the zonal structure of precipitation. I address the zonal structure sampling time by using the asymmetrical component growth of a given field between cumulative samples. Once the asymmetrical component average error of a zonal mean is smaller than the standard deviation between subsamples, we can be confident of the subsampling size. I also corroborate it with the similarity score with a threshold (95%). For most statistics and characteristics of the general circulation, 120 days is sufficient (Table A.2). For example, the ITCZ features

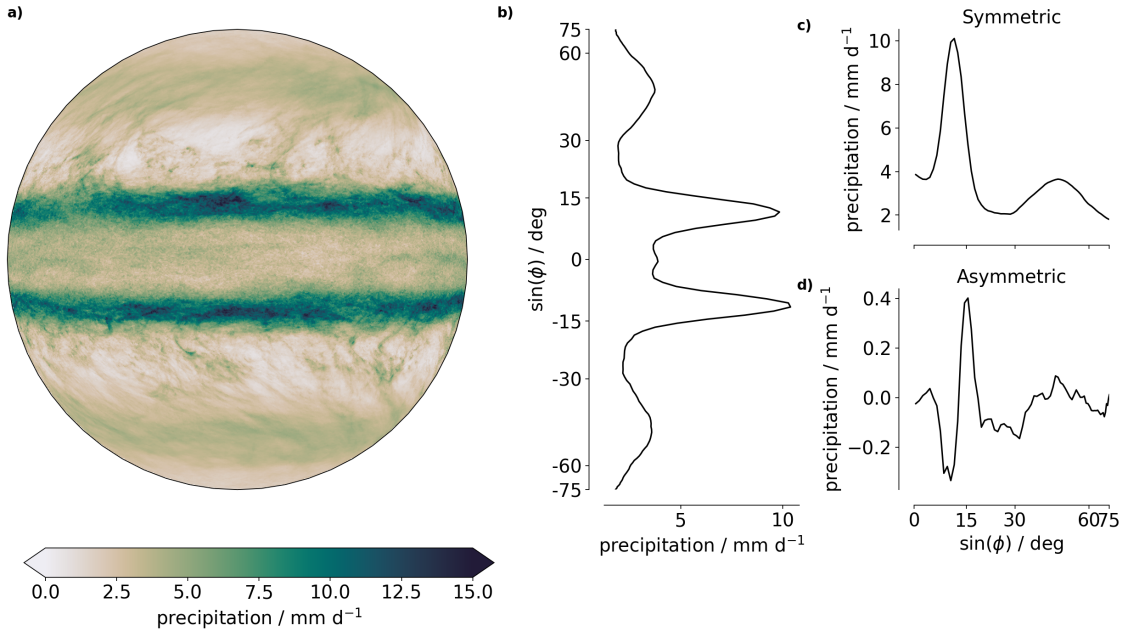


Figure 2.2: Example of the symmetrical characteristic of aquaplanet simulated with 5 km horizontal grid spacing over 90 days of integration time, showing precipitation as (a) global field, (b) zonal mean 75S-75N, (c) symmetrical component, and (d) asymmetrical component.

(latitude, intensity, and width) requires subsampling time around 20 days and sampling time up to 150 days, while the storm tracks (latitude and intensity) require a longer sampling time of 180 days (Figure A.2). For statistical quantities that do not achieve the required sampling time the asymmetrical component provides an estimate of uncertainty. Thus, the hemispheric asymmetry provides a means to quantify sampling errors due to under-sampling, which I use to evaluate convergence within uncertainty bands (section 2.2).

2.2 PHYSICAL CONVERGENCE

I developed a methodology based on the Richardson-extrapolation method incorporating asymmetrical components of our results to investigate convergence. The Richardson extrapolation method, widely used in the computational fluid dynamics (CFD) community (Phillips and Roy, 2014), uses lower-order discrete solutions to derive a higher-order approximation, enabling us to estimate the discretization error:

$$\varepsilon_h = \alpha_p h^p + \text{HOT} \approx \alpha_p h^p \tag{2.2}$$

where HOT describes high-order terms, h is the grid spacing, α_p a constant, and p the formal order of convergence, which depends on the discretization scheme used. Assuming that the numerical solution is in the asymptotic range of convergence, we can drop the HOT and approximate the discretization error of two systematically refined grids with asymptotic numerical solutions:

$$\varepsilon_h = \frac{u_{rh} - u_h}{r^p - 1} \tag{2.3}$$

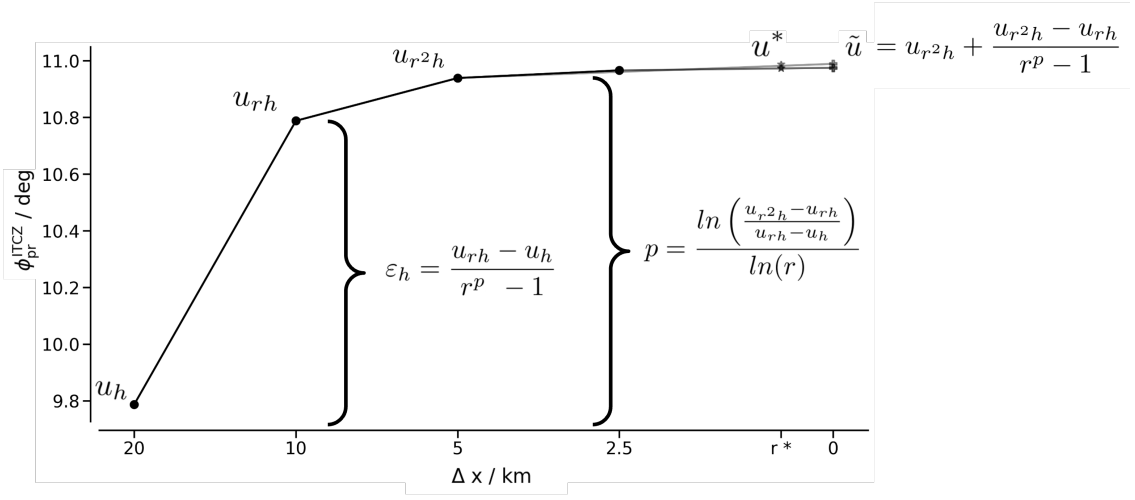


Figure 2.3: An example of the Richardson-Extrapolation method for ITCZ location ($\phi_{\text{pr}}^{\text{ITCZ}}$), which is defined by the latitude of the zonal mean of precipitation maxima, for horizontal grid spacing between 20 and 2.5 km, where the evaluated grid spacing [h, rh, and r^2h] are characterized by r, the grid refinement factor (in this case equal to 2), and the coarser horizontal grid spacing (h), p is the estimated convergence rate, u is the metric evaluated at each grid spacing, ε_h is the discretization error evaluated within h and rh, $\tilde{u}$ is the estimated asymptotic value using the evaluated grid spacing, and $r*$ is an estimated resolution to achieve an acceptable value within a certain desired accuracy ($u*$). Extended black and gray lines towards $\tilde{u}$ use the horizontal grid spacing of [10, 5, 2.5] km, and [20, 10, 5] km, respectively

where u is a characteristic or metric to evaluate between the two simulations with grid spacing h and rh respectively, with r the grid refinement factor and h is, therefore, the coarser grid spacing between the two evaluated experiments.

However, due to the complexity of the non-linear partial differential equations that compose the climate models and the numerical scheme used, we do not know the formal order of accuracy or convergence rate (p). We use an additional refined grid to estimate the convergence rate (r^2h):

$$p = \min\left(0.5, \frac{\ln\left(\frac{u_{r^2h} - u_{rh}}{u_{rh} - u_h}\right)}{\ln(r)}\right) \tag{2.4}$$

where 0.5 is a limiting factor to minimize unrealistic large error estimates (Phillips and Roy, 2014) that may emerge in non-convergent regimes. Once we estimate the convergence rate, we estimate the asymptotical value by:

$$\tilde{u} = u_{r^2h} + \frac{u_{r^2h} - u_{rh}}{r^p - 1} \tag{2.5}$$

Hence, our method requires three experiments with grid spacing [h, rh, and r^2h] to estimate convergence. Figure 2.3 displays the structure to evaluate the convergence of an asymptotical variable, the ideal case, where α_p is constant within the two consecutive discretization errors (ε_h and ε_{rh}). However, due to the non-linear nature of the climate and our estimated convergence rate, diverse metrics may display a non-smooth convergence pattern or become noisy. To solve this problem, I use a comparative approach, which

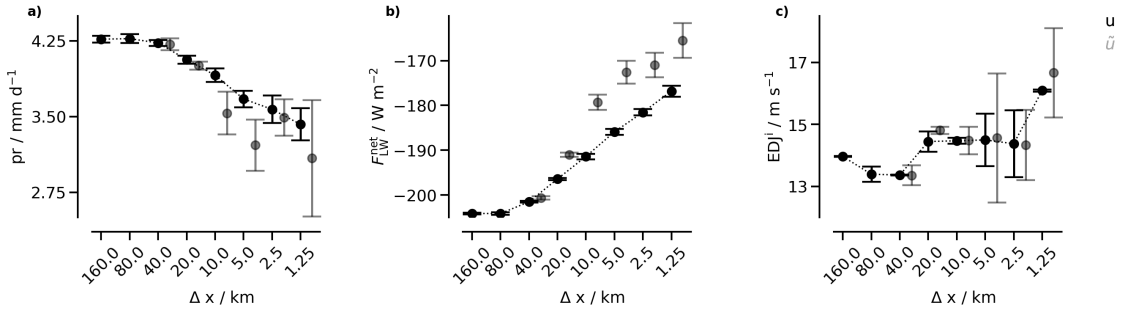


Figure 2.4: Examples of the convergence patterns of three different metrics using the convergence methodology developed (simplified analysis as it does not include the confidence metric, ζ): (a) reliable convergence pattern of global mean precipitation (pr) achieving convergence at 2.5 km horizontal resolution, (b) steady convergence pattern of global mean atmospheric net longwave radiation ($F_{\mathrm{LW}}^{\mathrm{net}}$) requiring horizontal resolution higher than 1.25 km to achieve convergence, and (c) unclear convergence pattern of Eddie-driven jet intensity EDJ^i showing an apparent convergence at a horizontal resolution of 10 km as it does not vary significantly with increasing resolution, but large uncertainty bars are present. In black, each metric (u) is evaluated at each horizontal grid spacing (x-axis), and in gray, each estimated asymptotical value ($\tilde{u}$) is assessed with the three previous horizontal grid spacing outcomes.

assesses the discretization error growth between two consecutive pairs and if it follows the discretization convergence rate ($\alpha_h = \varepsilon_h / h^p$), creating the metric:

$$\zeta = \frac{\varepsilon_h(p)}{\varepsilon_{rh}(p)} r^p \tag{2.6}$$

where $\zeta \approx 1$ indicates that we are in the asymptotical regime. Thus, I define a range of confidence based on ζ, where $|1 - \zeta| < 0.1$ is considered highly reliable, $\zeta < 0.25$ moderately reliable, $\zeta < 0.5$ questionable, and $\zeta > 0.5$ unreliable.

Additionally, undersampled metrics may enhance non-smooth convergence patterns and become more noisy. For example, Figure 2.4c displays oscillatory values at 2.5 and 1.25 horizontal grid spacing as the Eddie-driven jet intensity requires longer sampling times than 120 days to have robust statistics, which has not been achieved. To address uncertainties due to under-sampling, I evaluate each metric using the hemispherical symmetrical component $\hat{u}$ (doubling the sampling), and with its asymmetrical component δ_u, I analyze its uncertainty by re-calculating it within a plus and minus hemispherical asymmetrical component; obtaining an uncertain range $u = \hat{u} \pm \delta_u$. I use the functional approach (Hughes and Hase, 2012), corroborated with the Monte Carlo approach (Possolo and Iyer, 2017), to propagate the uncertainty δ_u through Eq. (2.5) and obtain the estimated asymptotic value ($\tilde{u}$) within its respective uncertainty range ($\delta_{\tilde{u}}$).

Figure 2.4 exemplifies a simplified version of convergence analysis for three different convergence patterns: reliable (a), steady (b), and unclear (c). A reliable and steady convergence pattern displays a steady convergence pattern as the horizontal resolution increases with their $\tilde{u}$ within moderate or greater confidence ($|1 - \zeta| < 0.25$). A reliable convergence pattern displays convergence within the horizontal grid spacing explored. For example (Figure 2.4a), mean global precipitation displays convergence at 2.5 km horizontal grid spacing, as the metric at 2.5 km and 1.25 km is within $\tilde{u}$ uncertainty bars. Meanwhile, steady convergence patterns do not achieve convergence even at 1.25 km and require finer horizontal grid spacing to achieve convergence. Figure 2.4b shows a

steady convergence pattern for global mean atmospheric net longwave radiation as, by increasing horizontal resolution, the difference within $\tilde{u}$ becomes smaller with a low convergence rate (p) of 0.5. In contrast, unclear convergence patterns display visual- or non-convergent patterns, i.e., regime change, with increasing horizontal grid spacing. Due to non-smooth convergence patterns, it display large uncertainty bars and low confidence ($|1 - \zeta| > 0.25$) in its $\tilde{u}$. For example (Figure 2.4c), the Eddie-driven jet intensity (EDJ^i) requires sampling times greater than 120 days to have robust statistics. At finer horizontal grid spacing than 5 km, it displays large uncertainty bars at each horizontal resolution and slight differences between them. This pattern leads $\tilde{u}$ to have large uncertainty bars and low confidence ($|1 - \zeta| > 0.25$). Moreover, as the 1.25 km horizontal grid spacing experiment has a short sampling time, EDJ^i is far from the 2.5 km horizontal resolution $\tilde{u}$. Additionally, I evaluate p for each metric, which estimates how fast the metric converges with increasing horizontal grid spacing. Using p and a desired error, i.e., standard deviation within ten-day sampling time, we can determine the horizontal grid spacing required to achieve convergence with Eq 2.2 for high confidence steady convergent metrics. However, it must be considered as an estimate.

3

KEY RESULTS

"Are there reasons to believe simulations if they do not converge? If models do not converge does it follow that they are incorrect?"

- David L. Williamson (Williamson, 2008)

Up to this point, we have laid out our methodology for assessing convergence and selected the model and configuration setup that captures the atmospheric climate's main characteristics, the global aquaplanet model with SST-QOBS. Its response to forcing, mimicking climate warming, is similar to simulations with more realistic configurations, i.e., AMIPs, and requires reduced computation resources to achieve the statistically steady state. In the sections that follow, we turn to the question posed by Williamson (2008), see above, and re-formulate it: By reducing the complexity of the model and parameterization that might depend on horizontal grid spacing, do we observe convergence? When refining horizontal grid spacing, does convergence reveal itself in a GSRM (ICON)? And does resolving small-scales and their interaction with large-scale reveal patterns of convergence in climate statistics? If so, does the response to forcing displays convergence?

To address these questions, I use the ICON with minimal parameterizations consistently for all experiments, except for shortening the time-step for higher horizontal grid spacings.

3.1 PHYSICAL CONVERGE IN A GSRM

In the first study (Appendix A), I performed aquaplanet experiments with horizontal grid spacing from 160 to 1.25 km. First, I explored the convergence of the global mean statistics, i.e., water and energy budget, representing the bulk properties of aquaplanet experiments. Subsequently, I characterized the large-scale circulation by the meridional overturning circulation, emphasizing the structure and intensity of the ITCZ, the boundaries of the tropics, and the storm tracks. Finally, I examined the zonal mean thermodynamic state in the tropics, focusing on the convergence of the temperature structure and cloud characteristics.

In accordance with previous studies (Kajikawa et al., 2016; Hohenegger et al., 2020), I observe a decrease in cloudiness with increasing horizontal grid spacing, specifically shallow clouds. With a systematic reduction of clouds, the atmospheric cloud radiative effect (ACRE) reduces, and a clear-sky radiative effect dominates the energy budget. The shortwave components of the energy budget are sensitive to horizontal resolution and display a steady and high-reliable convergence pattern but have yet to achieve convergence at kilometer-scale grid spacing. For example, net shortwave radiation at TOA ($F_{\mathrm{SW}}^{\mathrm{TOA}}$), displayed in Figure 3.1b, increases systematically as horizontal resolution increases, while their differences between asymptotical estimates becomes smaller. $F_{\mathrm{SW}}^{\mathrm{TOA}}$ is mainly controlled by the albedo effect of clouds, as there is no ice in aquaplanets. Precipitable water (prw) plays a crucial role in the radiative effect in clear-sky conditions,

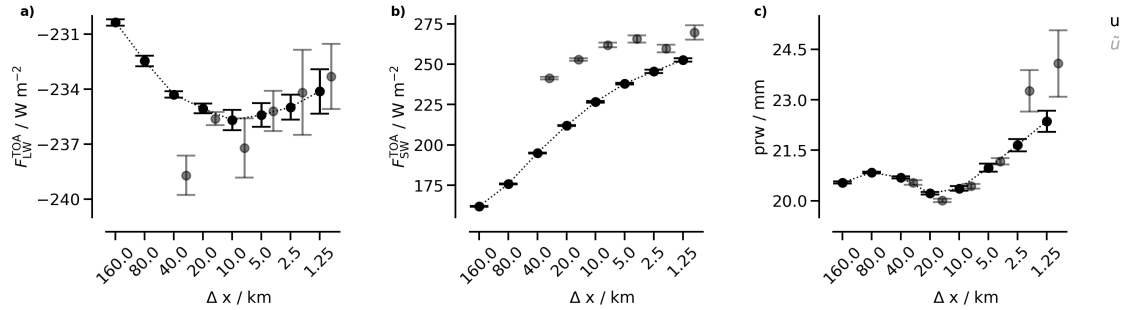


Figure 3.1: Simplified version of global mean statistics convergence analysis of (a) outgoing longwave radiation at TOA ($F_{\mathrm{LW}}^{\mathrm{TOA}}$), (b) net atmospheric shortwave radiation at TOA ($F_{\mathrm{SW}}^{\mathrm{TOA}}$), (c) precipitable water (prw). Similar to Figure 2.4

with effects on the outgoing longwave radiation at TOA ($F_{\mathrm{LW}}^{\mathrm{TOA}}$). With the increase of horizontal grid spacing, prw displays a minima at 20 km and an increase with finer horizontal resolution (Figure 3.1c). Interestingly, $F_{\mathrm{LW}}^{\mathrm{TOA}}$ shows a maximum at 10 km and then a reduction with a finer horizontal grid spacing (Figure 3.1a). I investigated this pattern further and observed that changes in the zonal circulation and, thus, in the prw zonal structure explain the mismatch between the prw minima and the $F_{\mathrm{LW}}^{\mathrm{TOA}}$ (not shown). Strikingly, only vertically integrated cloud ice content, precipitation, evaporation, and longwave cloud radiative effect at the top of the atmosphere display convergence at 2.5 km horizontal grid spacing (not shown).

Qualitatively, the meridional overturning circulation displays one single ITCZ at 160 and 80 km, an unstable ITCZ (transition from single to double) at 40 km, and a double ITCZ at higher horizontal grid spacing than 20 km (Figure 3.2). Rios-Berrios et al. (2020), using the Model for Prediction Across Scales-Atmosphere (MPAS-A, Skamarock et al., 2012) with cumulus parameterizations, also showed a double ITCZ within kilometer-scale horizontal grid spacing of up to 15 km. On the other hand, GCMs display different ITCZ structures with increasing horizontal resolution using the SST-QOBS (Landu et al., 2014) with slight different configurations (parameterizations).

Quantitatively, I characterize the ITCZ structure by its location (superscript ϕ), width (superscript w), and intensity (superscript i) as measured either in terms of precipitation (subscript pr), or the zonal meridional stream function at 500 hPa (subscript ψ). The ITCZ location, pr (Figure 3.2a) or ψ (Figure 3.2b), shifts polewardly with increasing horizontal grid spacing, and differences are below 0.5 deg within horizontal resolutions at 5 km. While the intensity reduces and displays slight differences within 5 and 2.5 km horizontal grid spacing. The ITCZ location and intensity display convergence at 2.5 km horizontal grid spacing (not shown). Metrics near the tropics boundary and extratropics present great variability and make the statistical convergence more challenging to quantify, as they require longer integration times than 120 days to have robust statistics. The Hadley cell edge ($\mathrm{ITCZ}_{\psi}^{\mathrm{w}}$), measured by the zero crossing of ψ, shifts polewards with increasing horizontal grid spacing (Figure 3.2b) with differences below 1 deg at finer horizontal resolution than 10 km. However, $\mathrm{ITCZ}_{\psi}^{\mathrm{w}}$ displays an unclear convergence pattern and sublinear convergence rate ($p < 1$, not shown).

In the extratropics, the storm tracks (ST), measured by the second precipitation maxima, shift polewards and become weaker with increasing horizontal resolution (Figure 3.2a). The storm track intensity (ST^{i}) and location (ST^{ϕ}) display apparent convergence at 5 km horizontal resolution with linear convergence rate and steady convergence, yet to converge at 2.5 km, with sublinear convergence rate, respectively (not shown). However,

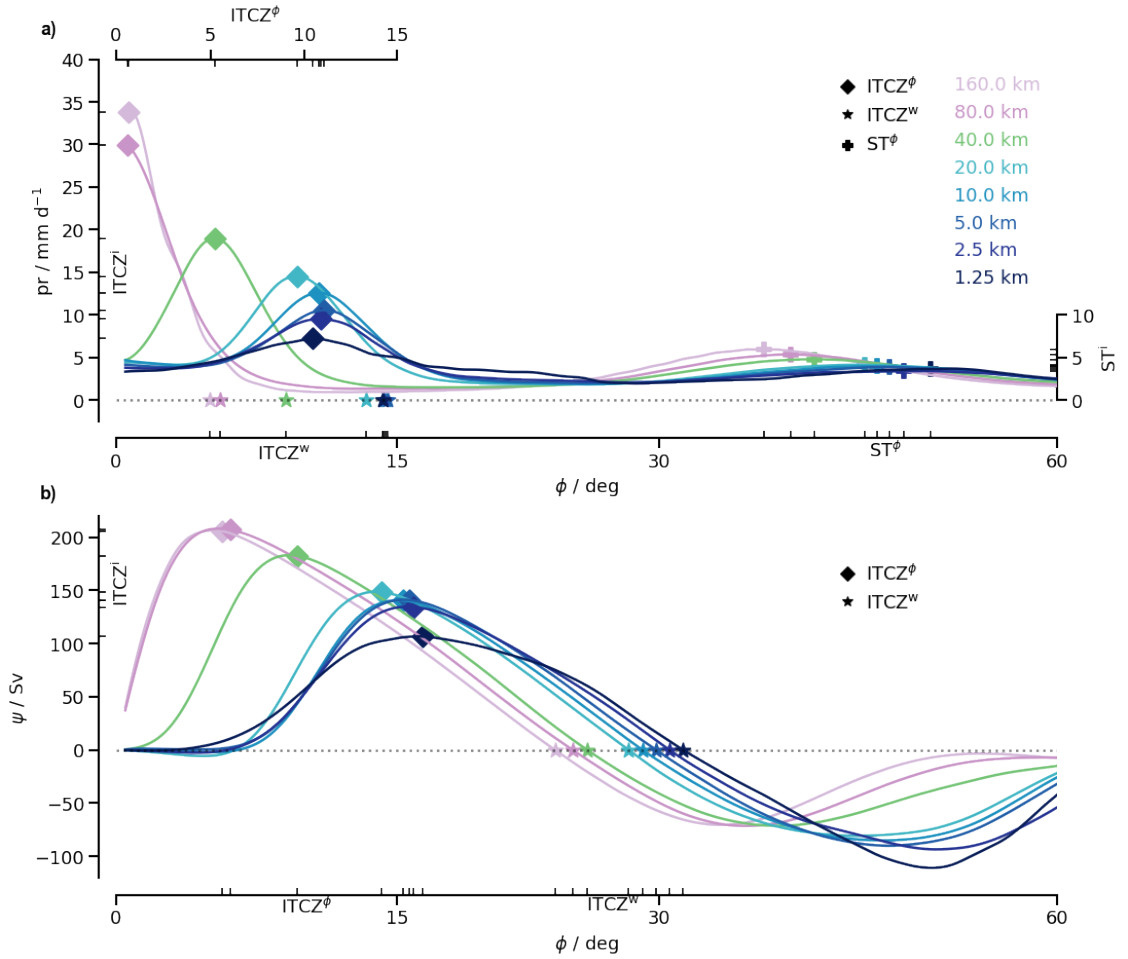


Figure 3.2: ITCZ zonal profile across horizontal resolutions (a): precipitation (pr), and (b): zonal-mean mass meridional stream function (ψ). The ITCZ, defined by its latitude and width, is marked with a diamond and star, respectively. The storm tracks latitude is marked with a cross for each horizontal grid spacing. On the x, y, and extra axes, I identify the ITCZ latitude (ITCZ$^{\phi}$), width (ITCZ^{w}), intensity(ITCZ^{i}), the storm tracks latitude (ST$^{\phi}$) and intensity (ST^{i}) for each horizontal grid spacing.

longer sampling times are needed to increase the confidence in the extratropical metrics convergence patterns.

The vertical thermal structure of the tropical atmosphere (Figure 3.3a) converges as I increase the horizontal grid spacing; it becomes less stable, and by comparing with the 2.5 km horizontal grid spacing experiment, the cold-point troposphere (CPT) warms quite remarkably at coarser experiments up to 20 km. CPT differences between horizontal resolutions finer than 20 km are less than 0.5 K and 200 m (in height), which are difficult to separate from variability.

The tropical boundary layer progressively dries (Figure 3.3b) and becomes less cloudy with a systematic reduction of shallow clouds (Figure 3.3c). The maxima of relative humidity in the boundary layer (RH_{max}) reduces up to 10 km horizontal grid spacing, where differences with finer horizontal resolutions are smaller than 0.5% in absolute value and 100 m in height. It displays convergence at 2.5 km horizontal grid spacing (not shown) with high confidence and linear convergence rate. Within the boundary layer, the maxima of cloud liquid water (clw_{max}) does not converge within the horizontal grid spacing studied; clw_{max} significantly reduces (Figure 3.3c) along with the maxima of

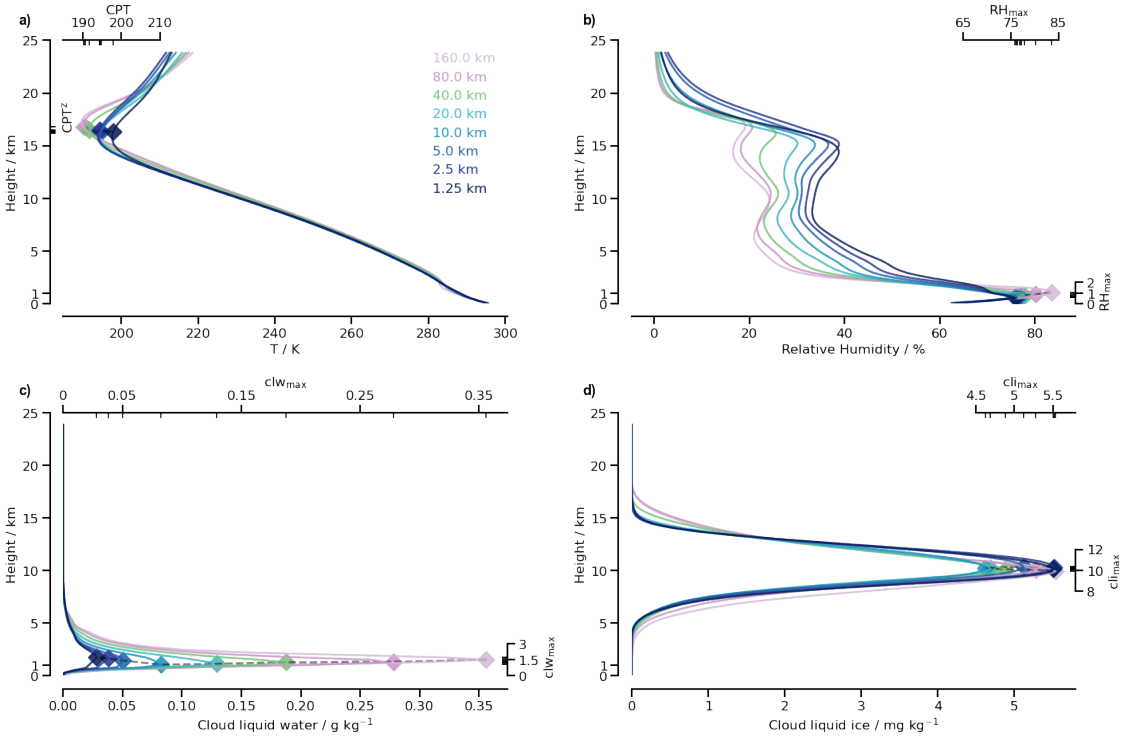


Figure 3.3: Tropical (30N to 30S) vertical structure with their respective characterization (metric) of (a) temperature characterized and the cold point tropopause (CPT), (b) relative humidity and maximum in the boundary layer (RH_{max}), (c) cloud liquid water content and its maximum (clw_{max}), and (d) cloud ice content and its maximum (cli_{max}). On the x, y, and extra axes, I identify each characterization (metric) magnitude and height for each horizontal grid spacing.

cloud cover in the boundary layer (not shown). Nevertheless, clw_{max} displays steady convergence with a sublinear convergence rate (not shown). Thus, shallow clouds are not sufficiently resolved within kilometer-scale horizontal grid spacing. Shallow clouds might not have a monotone convergence when horizontal grid spacing reaches decameter scales, thin stratiform clouds associated with the detrainment from trade-wind clouds will become better resolved and more pronounced, causing the planet to brighten (Stevens et al., 2001; Schulz and Stevens, 2023).

The tropical upper and middle troposphere (Figure 3.3b) become moister with horizontal grid spacing, and changes in the secondary peak in relative humidity are noticeable even at the finest horizontal resolutions. In the upper troposphere, ice clouds do not display monotonicity as I increase horizontal grid spacing (Figure 3.3d) and are subject to the meridional overturning circulation (single or double ITCZ). The convergence pattern of cloud ice maxima (cli_{max}) shows signs of convergence at 2.5 km, but the differences within the finest horizontal grid spacing in the order of 0.5 mg/kg generate large uncertainty and low confidence in the convergence analysis (not shown). Given the significant changes in circulation and humidity, cli_{max} are controlled mainly by microphysical and thermodynamical processes rather than resolution.

3.2 PHYSICAL CONVERGE OF CLIMATE RESPONSE TO +4K FORCING

In the second study (Appendix B), I extend the simulation of the first study (control), and using the same initial state, I performed forced simulations, mimicking climate

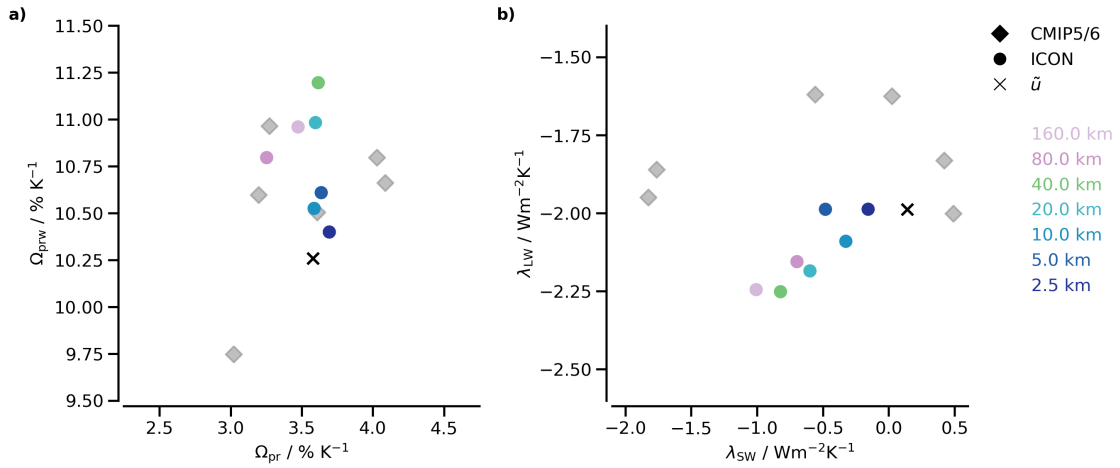


Figure 3.4: Global mean response to warming, (a) changes in the water budget, where the x-axis is the intensification of the hydrological cycle (Ω_{pr}) and y-axis the precipitable water increase (Ω_{prw}), and (b) changes in the energy budget, climate feedback (λ) components, shortwave (x-axis) and longwave (y-axis). Color circles are the result of our experiments from 160 to 2.5 km horizontal grid spacing, CMIP5/6 experiments in gray diamonds (as a reference), and the estimated convergence value for each metric ($\tilde{u}$), as a black x.

warming by increasing the SST by 4K uniformly for the same period. For the forced experiments, I discarded the first 45 days as the spin-up time. Experiments with 160 and 80 km, 40 to 5 km, and 2.5 km horizontal resolution have a simulation time of 360, 180, and 90 days, respectively. First, I explored the convergence of the energy balance at the top of the atmosphere and the intensification of hydrological cycle (precipitation) and precipitable water, representing the main response in climate change. Subsequently, I focus on the large-scale circulation structure, i.e., the contraction of the ascending branch of the Hadley cell. Finally, I examine the thermodynamical vertical response of the tropics, focusing on relative humidity and the clouds response.

The global energy balance at the top of the atmosphere shows a robust response as the horizontal grid is refined (Figure 3.4b). Agreeing with Retsch et al. (2019), the longwave climate feedback (λ_{LW}) dominates ECS. λ_{LW} becomes more positive with refinement and converges at 5 km horizontal grid spacing to a value of -2.1 W m^{-2} in accordance with spectral longwave feedback parameter estimated from satellite observations (Roemer et al., 2023). The shortwave climate feedback (λ_{SW}) displays high sensitivity to horizontal grid spacing, consistent with the convergence of the control experiments, becoming neutral at higher horizontal grid spacing with an oscillatory low-reliable convergence pattern yet to converge at kilometer-horizontal grid spacing (not shown). AMIP experiments at high horizontal grid spacing have observed similar patterns of λ_{SW} (i.e. Noda et al., 2019; Merlis et al., 2024b).

With warming, the precipitable water increases (Ω_{prw}) following the temperature dependence of saturation-specific humidity (Figure 3.4a), known as the Clausius-Clapeyron (CC) relationship (Trenberth et al., 2003). Similarly, precipitation increases, also refereed as the intensification of the hydrological cycle (Ω_{pr}), but at a slower rate, as is constrained by the change in atmospheric energy availability (i.e. Allen and Ingram, 2002; Held and Soden, 2006; Jeevanjee and Romps, 2018). For all horizontal grid resolution studied, the Ω_{pr} increase rate is slower than Ω_{prw} and have different convergence patterns. The Ω_{pr} displays apparent convergence at 20 km horizontal grid

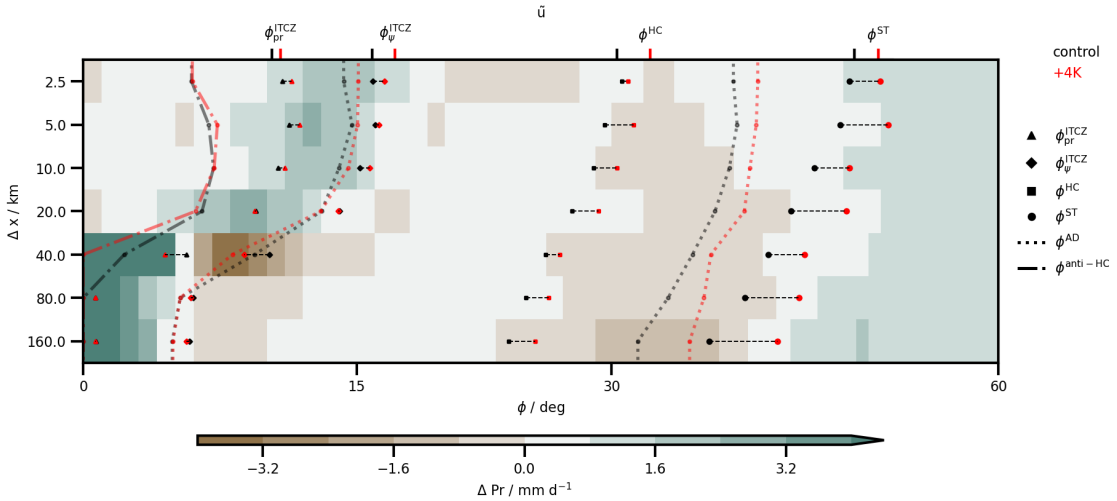


Figure 3.5: The response pattern of the large-scale circulation structure, represented by the intertropical convergence zone (ITCZ), in terms of precipitation change in response to uniform warming of + 4K, from 160 to 2.5 km horizontal grid spacing (y-axis). I define the ITCZ structure in terms of its location (maxima of precipitation: $\phi_{\mathrm{pr}}^{\mathrm{ITCZ}}$; and maxima of the stream function at 500 hPa: $\phi_{\psi}^{\mathrm{ITCZ}}$), Hadley cell edge ($\phi^{\mathrm{HC}}$), storm track location (second maxima of precipitation; ϕ^{ST}), atmospheric desert (ϕ^{AD}), and anti-Hadley cell edge ($\phi^{\mathrm{anti-HC}}$).

spacing converging at 5 km towards 3.6 %K^{-1} (not shown), higher than the estimated 2-3 %K^{-1} from previous studies (Allen and Ingram, 2002; Held and Soden, 2006). In contrast, Ω_{prw} does not converge at kilometer-scale horizontal grid spacing and displays a steady convergence pattern towards 10.25 %K^{-1} (not shown), higher than the CC relationship of 7 %K^{-1} and in line with CMIP5 aquaplanet response (Medeiros et al., 2015).

Changes in the hydrological cycle and the meridional overturning circulation are closely interconnected (Chahine, 1992; Stevens and Bony, 2013a). However, the response of the overturning circulation, i.e., changes of the ITCZ location, displays different responses between models (i.e., Medeiros et al., 2015). Figure 3.5 displays the zonal response of precipitation (forced-control), the location of the ITCZ measured by precipitation ($\phi_{\mathrm{pr}}^{\mathrm{ITCZ}}$) and the stream function ($\phi_{\psi}^{\mathrm{ITCZ}}$), the atmospheric dessert (ϕ^{AD}) start and end (the zero crossing of the water budget), the Hadley cell edge (ϕ^{HC} - the zero crossing of the stream function), the anti-Hadley cell edge ($\phi^{\mathrm{anti-HC}}$ - the crossing zero near the equator of the stream function, Adam, 2021), and the storm track location (ϕ^{ST} - the second maxima of precipitation). Regions with high precipitation (wet regions) and dry regions show an increase and decrease in precipitation, respectively, following the paradigm of the "rich-get-richer mechanism" across horizontal resolutions (Chou and Neelin, 2004; Chou et al., 2009).

In accordance with previous studies (i.e., Liu et al., 2012; Vallis et al., 2015; Son et al., 2018), the Hadley cell weakens and expands (poleward and upwards), and the atmospheric desert expands for all horizontal grid spacing investigated in response to warming (Figure 3.5). ϕ^{HC} shifts polewards robustly and displays unclear apparent visual convergence at 5 km horizontal grid spacing towards an estimated convergence value of 1.6 deg (not shown). Along the poleward expansion of the Hadley cell, the storm tracks shift poleward. ϕ^{ST} displays a similar convergence pattern to that of ϕ^{HC} with an estimated convergence value of 1.9 deg (not shown). However, the great variability

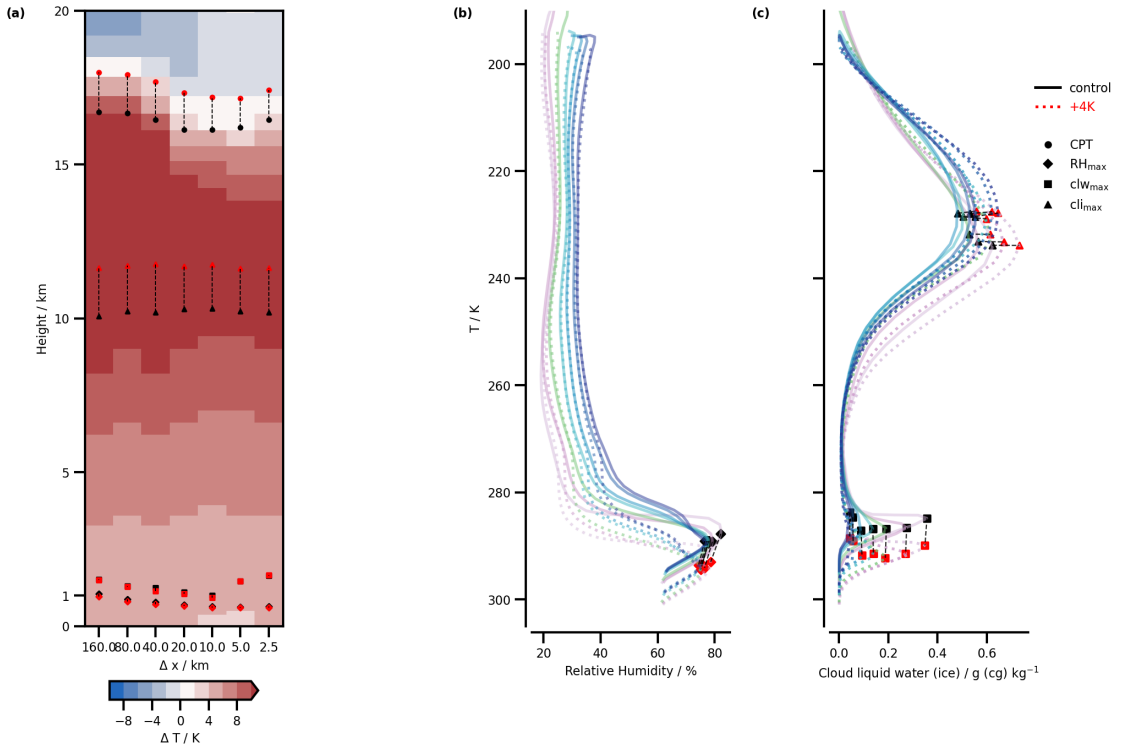


Figure 3.6: Thermodynamic vertical response in the tropics of (a) temperature and (b) relative humidity in model level, and (c) cloud ice and liquid water in temperature space from 160 to 2.5 km horizontal grid spacing. I characterize the thermodynamic vertical structure regarding the cold point tropopause (CPT) and the maxima of relative humidity (RH_{max}), cloud liquid water (clw_{max}), and cloud ice cli_{max}.

in the limits of the Hadley cell and extratropics makes the statistical convergence more challenging to quantify and requires longer sampling times for a robust estimate.

In the tropics, specifically the ascending branch of the Hadley cell, I observe distinct sensitivity to horizontal grid resolution and associated with the ITCZ structure in control simulations. Byrne et al. (2018) defined the ITCZ width as the width of the ascending branch of the Hadley cell, which can be defined as the $\phi_{\psi}^{\mathrm{ITCZ}}$ if there is no anti-Hadley cell. If an experiment has an anti-Hadley cell, I define the width as the distance between the edge of $\phi^{\mathrm{anti-HC}}$ and $\phi_{\psi}^{\mathrm{ITCZ}}$. The ITCZ location does not shift in experiments with 80 and 160 km horizontal resolution, while the ITCZ width displays a contraction with absolute value smaller than 0.2 deg. When the control simulation has an unstable ITCZ (40 km horizontal grid spacing), the ITCZ location shifts equatorward, and the ITCZ width expands in response to warming. At finer horizontal grid spacing, the ITCZ location shift polewards, and the ITCZ width expands by an absolute value smaller than 1 deg; differences between horizontal grid resolutions are small and make the estimate unreliable (not shown). Interestingly, the ITCZ width displays two modes in our experiments, contrary to previous studies (i.e., Byrne and Schneider, 2016; Watt-Meyer and Frierson, 2019), where the ITCZ contracts whenever there is a single or double ITCZ in control experiments.

Under global warming, the tropopause height and stability are expected to increase as the outgoing longwave radiation is constrained, on average, to the incoming solar radiation, and the moist adiabatic lapse rate decreases, especially in the tropics (Vallis et al., 2015). Across the experiments, the cold point tropopause (CPT) and maxima

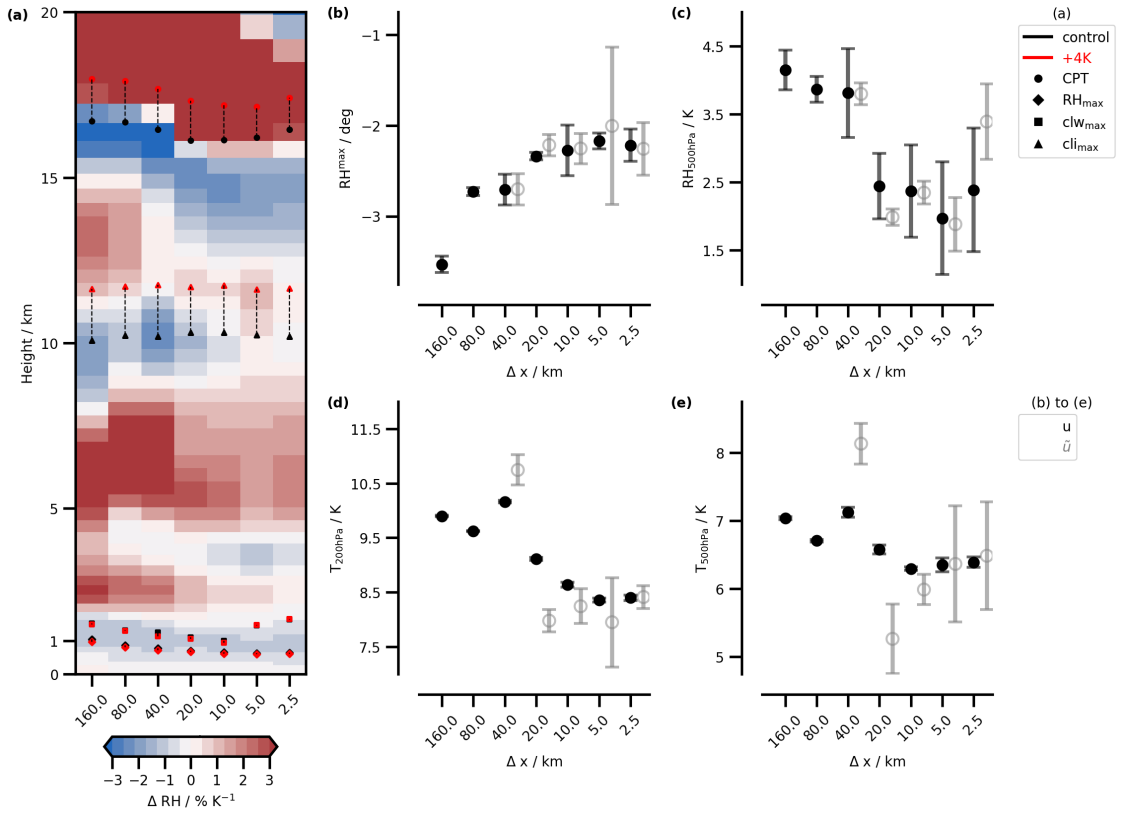


Figure 3.7: Thermodynamic vertical response in the tropics of (a) relative humidity in model level, similar to Figure 3.6, and simplified version of the statistics convergence analysis of the response to warming of (b) relative humidity maxima in the boundary layer (RH^{max}), (c) relative humidity at 500 hPa (RH_{500hPa}), (d) temperature at 200 hPa (T_{200hPa}), and (e) temperature at 500 hPa (T_{200hPa}), similar to Figure 2.4

of cloud ice have a robust upward movement (1 km) as response to warming with differences between experiments of 100 m (Figure 3.6a), which are difficult to separate from variability. CPT increases by 2 K for horizontal grids spacing finer than 20 km and differences between experiments are below 0.25 K. Likewise, the vertical temperature profile displays significant warming. The upper levels of the troposphere (200 hPa ≈ 12 km) warm by about 8.4 K (Figure 3.7d), and the middle levels of the troposphere (500 hPa ≈ 6 km) warm comparatively less by 6.5 K (Figure 3.7e). The warming in the troposphere, measured at 200 hPa and 500 hPa displays apparent convergence at 5 km horizontal grid spacing with large uncertainty bars due to small differences within higher horizontal resolutions. CPT change and the temperature response at 200 and 500 hPa agrees with recent studies (Guendelman et al., 2024).

It is well-known that climate models' response to warming tends to maintain a near-constant relative humidity (i.e., Manabe and Bryan, 1969; Held and Soden, 2006), which is supported by observations (Douville et al., 2022). Jeevanjee and Romps (2018) noted that relative humidity above 1 km in the free troposphere is invariant in temperature coordinates and does not depend on warming. Figure 3.6b shows the change in the vertical structure of relative humidity in temperature coordinates and Figure 3.7a in model levels. It becomes evident that the boundary layer dries while the middle and upper troposphere becomes moister in a warmer climate. In the boundary layer, the maxima of relative humidity decreases by 2.3 % with reliable convergence patterns at 10 km horizontal grid spacing (Figure 3.7b). The increase of water vapor and relative

humidity greatly impact the longwave climate feedback as the emission height increases. At 500 hPa, approximately the emission height, I observe an increase of relative humidity with apparent visual convergence at 10 km horizontal grid spacing, but with unclear confidence and significant uncertainty bars towards 2.4 % (Figure 3.7c), increasing the emission temperature by 2.4 K (not shown).

The response of clouds to warming remains uncertain and is considered the major source of uncertainty of climate sensitivity estimates from climate models in CMIP6 (Zelinka et al., 2022). However, there are robust responses in models, such as the anvil top cloud temperature and height. According to the hypothesis of the fixed anvil temperature (FAT, Hartmann and Larson, 2002) and the proportionally higher cloud base warming (PHAT, Zelinka and Hartmann, 2010) hypothesis, anvil top cloud temperature should remain almost constant. I test this hypothesis across experiments by changes in the cloud top temperature (not shown) and temperature changes of the maxima of cloud liquid ice (Figure 3.6c). The cloud top in our experiments has a slight increase of 1 K for all resolutions (not shown), thus, agreeing with the PHAT hypothesis. Following the PHAT hypothesis and the increase of stability in upper levels of the troposphere, Bony et al. (2016) suggested a contraction of anvil clouds (reduction of top cloud cover) in the deep tropics (Iris effect, Lindzen et al., 2001). However, the Iris effect is not observed clearly across the different horizontal grid spacing explored. Instead, a slight increase of anvil clouds, in line with NICAM simulations (Noda et al., 2019) is observed. Meanwhile, shallow cloud amounts do not show significant changes across resolutions (Figure 3.6c).

4 CONCLUSION

> *"...If a model converges it is not necessarily correct. However, one would have more confidence in the results if multiple schemes converged to the same statistical states..."*
>
> - David L. Williamson (Williamson, 2008)

This dissertation ventures into the convergence in climate response to warming using a Global Storm Resolving Model by developing a new methodology to quantify convergence in a simulated climate and its response to warming. At its core, we aim to answer the question: *With the increase of horizontal grid spacing and resolving small-scale processes (e.g., convection), does the large-scale climate in the present day display a qualitative and quantitative convergence? If so, does the response to warming display a qualitative and quantitative convergence, too?* While GSRMs do not solve all modeling challenges, they provide a more physical framework by solving a consistent set of equation over a broader range of scales (Satoh et al., 2019). By resolving more physics (including fluid-dynamic and non-fluid-dynamical processes) and minimizing the use of complex, resolution-dependent parameterizations, GSRMs reduce the entanglement of parameterization and numerical errors. Chapter 2 outlines the advantages of aquaplanet experiments with reduced computational resources to achieve robust statistics and the developed methodology to evaluate convergence, with an adapted Richardson-Extrapolation method. This work provides a way to examine GSRMs from the point of view of convergence. Chapter 3 presents the main findings of this dissertation by using the climate model ICON retaining only parameterization for the radiant energy transfer, a single moment bulk microphysical parameterization, and 3D turbulent mixing.

4.1 ANSWERING THE RESEARCH QUESTIONS

Let us examine the research questions outlined in Chapter 1.2.

RQ1. Does the GSRM ICON display convergence with increasing horizontal grid spacing?

To answer this question, I developed a methodology to evaluate convergence with a sampling time of at least 120 days based on the Richardson-Extrapolation method. I integrated the symmetrical characteristics of aquaplanets in the methodology to derive confidence estimates. To test convergence, I performed a series of aquaplanet experiments using ICON with horizontal grid spacing from 160 to 1.25 km.

Qualitatively speaking, the large-scale circulation structure, measured by the ITCZ structure, transitions from single to double at 40 km horizontal grid spacing. At 10 km horizontal grid spacing, the large-scale circulation structure displays convergence. The main features of the ITCZ (location and intensity) show reliable convergence at a horizontal grid spacing of 1-5 km (Figure 3.2). However, features in the subtropics, i.e.,

Hadley cell edge, and in the extratropics, show apparent convergence at a horizontal grid spacing of 1-5 km as their variability makes statistical convergence more challenging to diagnose in short simulations (Figure 3.2). The vertical structure in the tropics, measured by temperature, relative humidity, and cloud ice, displays convergence with small differences between horizontal grid resolutions in the range from 5 km to 1.25 km (Figure 3.3). The boundary layer dries with increasing horizontal grid spacing and converges faster than the middle and upper troposphere. The middle and upper troposphere becomes warmer and moister with increasing horizontal resolution and displays changes by a few percent in relative humidity. However, the cloud liquid water maxima are sensitive to the horizontal grid spacing and do not show convergence at kilometer-scale horizontal grid spacing (Figure 3.3). Thus, shallow clouds are underresolved at kilometer-scale horizontal grid spacing, affecting the convergence of the shortwave component of the energy budget. The longwave components of the energy budget, i.e., outgoing longwave radiation, converge with increasing horizontal resolution and achieve convergence at 2.5 km horizontal resolution (not shown).

RQ2. If a GSRM shows convergence with increasing horizontal grid spacing, does its response to warming also show convergence

To answer this question, I extended the aquaplanet experiments in the first study, refereed to as "control", and by using the same starting initial state, I performed "forced" experiments, mimicking climate warming with a uniform increase of SST by 4 K, for horizontal grid spacing from 160 km to 2.5 km. I focus on different aspects of global and regional responses to climate change and use the convergence methodology developed in Chapter 2 to assess convergence.

Overall, the response to global warming in aquaplanet experiments shows convergence as we increase the horizontal grid spacing. From the global perspective, canonical metrics, like the hydrological cycle intensification, converge at kilometer-scale horizontal grid spacing, except for shortwave climate feedback and precipitable water increase, similar to "control" experiments. In the large-scale structure, the mean meridional circulation response, i.e., ITCZ location shift, converges at kilometer-scale horizontal grid spacing.

The overturning circulation displays three different response patterns in the location and width of the ITCZ relative to the "control's" climate structure converging at kilometer-scale grid spacing (Figure 3.5):

a) no movement and contraction (single ITCZ) at 160 and 80 km

b) equatorward shift and expansion (unstable ITCZ) at 40 km

c) poleward shift and expansion (double ITCZ) for horizontal grid spacing finer than 20 km

The Hadley cell expands for all horizontal grid spacings. The Hadley cell edge and storm track location have a poleward shift with apparent convergence at 5 km horizontal grid spacing. However, their great variability makes statistical convergence more challenging to diagnose. Similar to the "control" climate, the tropical boundary layer dries with horizontal grid refinement, while the tropical free troposphere becomes moister and warmer, which modifies the emission temperature in clear-sky conditions (Figure 3.6).

The cloud top temperature change follows the hypothesis of the fixed anvil temperature (FAT, Hartmann and Larson, 2002) and the proportionally higher cloud base

warming (PHAT, Zelinka and Hartmann, 2010) hypothesis with a slight temperature increase (1K) as a response to warming (Figure 3.6). Cloud radiative effects, longwave and shortwave, compensate for each other in the tropics. However, we observe significant changes in the extratropics, where mixed-phased clouds are more common, with an increase of cirrus and shallow clouds. Bony et al. (2016) estimated an increase of anvil clouds (Iris effect, Lindzen et al., 2001) based on the thermodynamic equation of convective mass fluxes and the usage of radiative-convective equilibrium (RCE) experiments within three GCMs. However, within our experiments, there is no robust decrease of anvil clouds. This agrees with previous studies of global climate change using NICAM (Noda et al., 2019). Jeevanjee (2022) explored the possibility that the convective max flux closure (Bony et al., 2016) is only partially accurate since relative humidity and precipitation efficiency might not remain constant, which is observed in our experiments (Figure 3.6). Meanwhile, in the subtropics, there is no robust decrease of shallow clouds, which is expected (i.e., Rieck et al., 2012). Instead, there is an increase in cirrus clouds.

4.2 FINAL REMARKS AND OUTLOOK

We do not know the idealized aquaplanet experiments' theoretical solution or climate as it cannot be solved analytically. We could aim to make observations, but the closest planet with a probable liquid surface is in the planetary system Kepler-138, which is just about 218 light-years distant and might be still very different to the idealized aquaplanet. Nevertheless, by increasing the horizontal grid spacing and avoiding complex parameterization, we have shown that a GSMR (ICON) does show convergence in the large-scale circulation structure and displays a robust response under climate warming.

Although the developed methodology offers valuable insights into climate models, it has caveats. Metrics with values near zero or small differences between neighboring horizontal resolution experiments might exhibit oscillatory or unstable convergence rates. Thus, the confidence in their analysis is low. In cases where visual convergence is apparent, we can be confident of convergence to a certain extent. As highlighted by Williamson's (2008) quote, the observed convergence increases our confidence in estimates of climate change and atmospheric processes in GSRMs, particularly in ICON. However, convergence with increased horizontal resolution does not guarantee the correctness of the climate model. In GSRMs, non-fluid dynamic component parameters or numerical representation of physical processes might influence the statistical climate; hence, how should we interpret the observed convergence?

The most obvious statement is that a model with kilometer-scale horizontal grid spacing can represent numerically the observed convergent metrics at those horizontal resolutions. Nevertheless, the estimated value of convergent metrics at kilometer-scale might be different according to the parameters within the remaining parameterization, i.e., mixing length scale in the 3D turbulence scheme. What remains to be answered is how large the impact of parameters in the remaining parameterization is on the climate state and in climate's response to warming. To test this question, I ran a set of experiments at 20 km horizontal grid spacing varying cloud liquid water conversion to precipitation, the horizontal diffusion and mixing length scale within the 3D turbulence mixing scheme, minimum wind threshold for surface evaporation, horizontal and vertical water vapor diffusion, and the numerical sub-time step of the dynamical core and horizontal numerical diffusion to test the first part of the question (Figure 4.1). By changing the numerical diffusion by a factor of 1024, the maxima of precipitation displays the largest

shift, but changes in cloud liquid water and ice are minimal (Figure 4.1a). Changes in microphysics, cloud liquid water conversion to precipitation, displays changes in the middle troposphere in the vertical structure of the tropics (Figure 4.1b). However, cloud ice does not show significant changes. Changes in the turbulence scheme displays an increase in cloud liquid water and a decrease in cloud ice, and changes in sub-time steps do not show a significant change overall. Nevertheless, despite strongly changing the diverse parameters, there is no substantial change in the large-scale meridional circulation structure, as precipitation maxima do not shift significantly (within 1-2 deg). These changes are small compared with changes due to an increase of horizontal grid resolution. From these results, I conclude, that changes in the large-scale meridional circulation structure and convergence are robust regardless of parameter changes. However, the interplay between different parameters within parameterizations might enhance the shift in circulation, but it would require large changes in the parameterizations. The question of how the parameter changes will affect the climate response under warming merits further investigation.

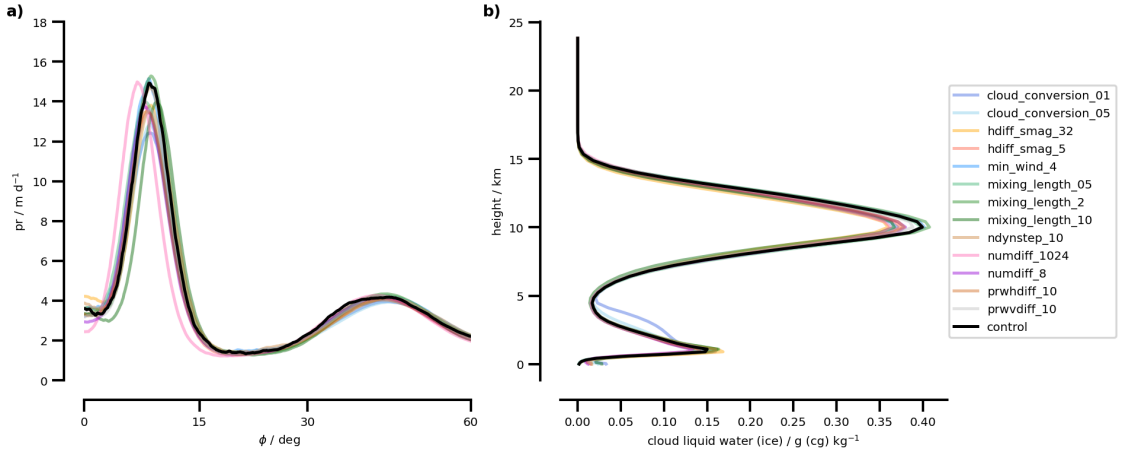


Figure 4.1: A sensitivity study varying parameters in diverse paramterization, individually, using horizontal grid spacing of 20 km. a) Zonal structure of precipitation, b) tropical vertical structure of cloud liquid water and cloud ice content. The parameters modified are `cloud_conversion`, cloud liquid water to precipitation with reducing factors of [0.1, 0.5], `hdiff_smag`, horizontal diffusion of water vapor in the turbulence scheme with increasing factors of [5 , 32], `min_wind`, minimum wind threshold for evaporation changed from 1 m/s to 4 m/s, `mixing_length`, mixing length scale in the turbulence scheme varying factors of [0.5, 2, 10], `ndynstep`, increase the numerical sub-time step in the dynamical core from 4 0 10, `numdiff`, horizontal numerical diffusion coefficent increase by a factor of [8, 1024], `prwhdiff`, horizontal water vapor diffusion increase by a factor of 10, and `prhdiff`, vertical water vapor diffusion increase by a factor of 10.

One might also be tempted to argue whether the convergence rate with horizontal grid refinement depends on parameters within parameterizations or the dynamical core of the GSRM. I compared two different ICON versions by running a new set of experiments using the latest version of ICON (hash 4dd46bd54) with sampling time of 360 days. In the new version, the dynamical core of ICON changed from conserving buoyancy to conserving energy. Figure 4.2a displays the zonal precipitation for both ICON versions with a horizontal grid spacing of 20, 10, and 5 km. Both versions of ICON display the same convergence trend in the evolution of the zonal precipitation, achieving only slight different climate states. The tropical vertical structure (Figure 4.2b) displays differences

in the relative humidity near the boundary layer between the two ICON versions, but the differences decrease with increasing horizontal grid spacing and differences between horizontal grid resolutions are much larger. Thus, ICON display similar convergence rate with changes in the dynamical core, hence, changes in the dynamical core have only a small impact on the climate state.

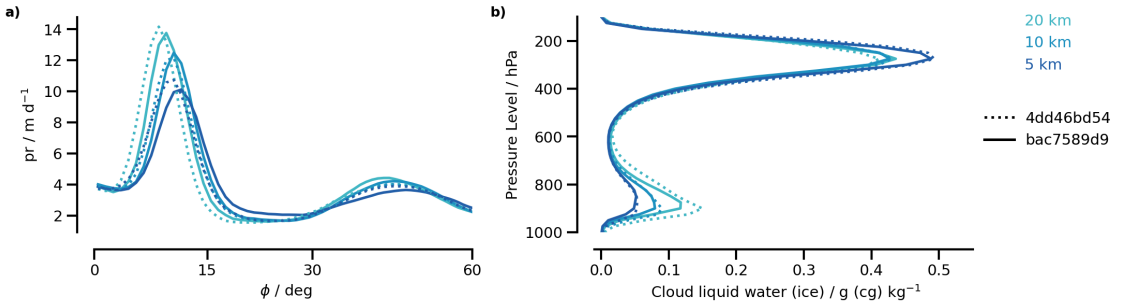


Figure 4.2: A sensitivity study using two different version of ICON with changes in the dynamical core, from buoyancy (bac7589d9) to energy conservative (4dd46bd54). a) Zonal structure of precipitation, b) tropical vertical structure of relative humidity with horizontal grid spacing from 20 to 5 km.

In this thesis, I focused on convergence in the tropics and the large-scale circulation. Figure 3.2 shows that extratropical metrics, i.e., the storm track location (ST_ϕ) and the Hadley cell edge (HC_ϕ), do not show convergence within the horizontal grid spacing studied. In contrast, Lu et al. (2015) suggested that the midlatitude metrics converge already at 50 km horizontal grid spacing. This discrepancy raises the question of what drives convergence in the midlatitudes metrics in "control" climate and its response to warming.

Lu et al. (2007) used observation to explain that the expansion of the Hadley cell as a response to warming is driven by the increase in the subtropical stability, which shifted the baroclinic instability zone poleward, hence HC_ϕ shifts. It is known that changes in ϕ^{HC} have effects in ST_ϕ (i.e., Kang and Polvani, 2011; Ceppi and Hartmann, 2013). However, HC_ϕ is determined by a complex interaction between the tropical circulation strength (meridional energy flux) and Eddie-mean flow adjustment. Changes in the deep tropics, i.e., convective stability, lead to an expansion or contraction of the Hadley cell, pushing the baroclinic instability zone (Mbengue and Schneider, 2013; Watt-Meyer and Frierson, 2019) polewards. At the same time, changes in the extra-tropical SST gradients plays a role in the distance between the ST_ϕ and HC_ϕ through transient eddie energy flux variations, which is affected by diffusivity (Mbengue and Schneider, 2018). Figure 4.3a displays the shift of ST_ϕ versus changes in $ST_\phi - HC_\phi$, representing changes by local processes (eddie energy flux variations). As we increase horizontal grid spacing, $ST_\phi - HC_\phi$ and ST_ϕ de-correlate. In contrast, the correlation between HC_ϕ and ST_ϕ (non-local processes) does not change with increasing horizontal grid spacing. Hence, a question arises whether uncertainties and biases in the simulated tropical climate affect the convergence of the storm tracks and therefore the response to warming and the representation of the mid-latitudes.

In unison, shallow marine boundary layer clouds do not converge at the horizontal grid spacing studied. Shallow marine boundary layer clouds convergence requires horizontal resolution in the order of hectometer to converge, consistent with evidence from large-eddy simulations (LES, i.e., Stevens et al., 2020). Global LES experiments could be a relevant tool to investigate global warming, but the computational expenses to

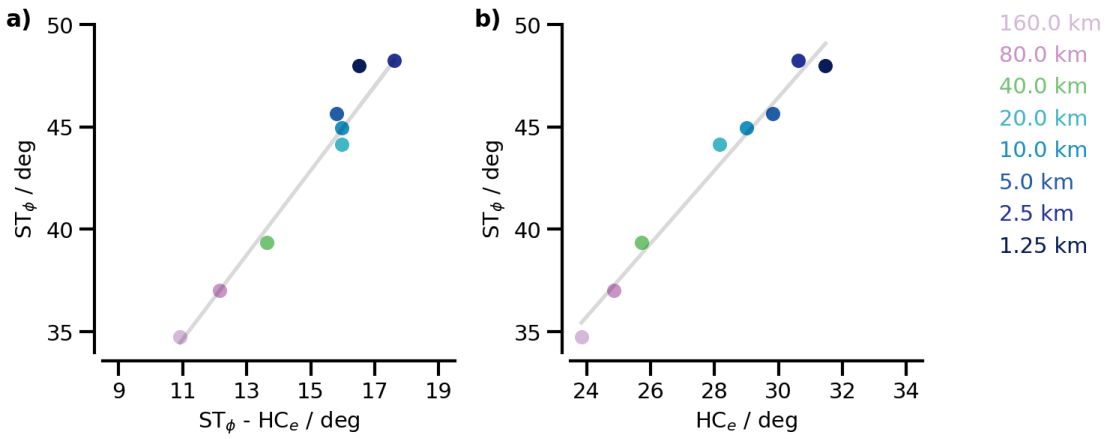


Figure 4.3: Extra-tropics convergence based on the Hadley cell edge (crossing zero of the stream function; ϕ^{HC}) and storm track location (second maxima of precipitation; ϕ^{ST}). a) Local driver estimate measured by $\phi^{HC} - \phi^{ST}$ increase (x-axis) and ϕ^{ST} shift (y-axis), b) non-local driver estimate measured by the ϕ^{HC} expansion (x-axis) and ϕ^{ST} shift (y-axis) with increasing horizontal resolution form 160 km to 1.25 km for control experiments.

obtain robust statistics would be challenging with today's supercomputers. We could be tempted to use shallow convection parameterizations to tame shallow marine boundary layer clouds within kilometer-scale horizontal grid resolution (i.e., Bogenschutz and Krueger, 2013; Rio et al., 2019), as done in other GSRMs (i.e., Bechtold et al., 2014; Chinita et al., 2023). However, this path would require a re-evaluation of such parameterizations and the uncertainity that might arise in climate change response by using them.

As different GSRMs use a range of grids with different numerics, topology, and parameterization suits, it is not clear if one GSRM converges another GSRM with similar design philosophy would converge too. While ICON has displayed convergence with increasing horizontal resolution, the question of whether other GSRMs show a similar convergence pattern lingers. Conversely, do other GSRMs arrive at a similar statistical climate state using the same aquaplanet configuration? Recently, (Rios-Berrios et al., 2022) performed a series of SST-QOBS aquaplanet experiments with the Model for Prediction Across Scales-Atmosphere (MPAS-A, Skamarock et al., 2012), keeping cumulus parameterization, from 120 to 15 km horizontal grid spacing and observed a double ITCZ with apparent convergence in its structure. Interestingly, they performed an extra aquaplanet experiment deactivating cumulus parameterization within a band of 20S-20N. It displayed a different climate state, which could signal the non-convergent behavior of the climate model or that the different experiments are not comparable to each other. To answer those questions within a common framework, Tamaki Suematsu, Daisuke Takasuka and I have proposed aquaplanets experiments, denominated Im(DYAMOND3), as part of the third phase of DYAMOND (Takasuka et al., 2024).

Overall, as horizontal resolution increases, the response to warming is robust and displays convergence. However, how does it compare with other models? I compare the response of aquaplanet experiments to recent results of AMIP simulations within the GSRM community, as the response of aquaplanet experiments are similar (Medeiros et al., 2015), and compare with observations. The muted shortwave climate feedback, hydrological cycle intensification over the ocean, and warming in the tropical vertical structure agree with recent GSRM studies (Noda et al., 2019; Guendelman et al., 2024;

Merlis et al., 2024a,b). Meanwhile, the longwave climate feedback is in agreement with spectral estimates from satellite observations(Roemer et al., 2023). Thus, we are more confident that a better understanding and estimations of climate response to warming lie within kilometer-scale horizontal grid spacing.

Part II

APPENDICES

A

APPENDIX A

The work in this appendix is in revision for the *Journal of Advances in Modeling Earth Systems (JAMES)* as

Peinado-Bravo, A., Klocke, D., & Stevens, B. “Horizontal grid spacing convergence of aquaplanet experiments."

Horizontal grid spacing convergence of aquaplanet experiments.

Angel Peinado-Bravo[1,2], Daniel Klocke[1],Bjorn Stevens[1]

[1] Max Planck Institute for Meteorology, Hamburg, Germany
[2] International Max Planck Research School on Earth System Modelling, Max Planck Institute for Meteorology, Hamburg, Germany

Manuscript submitted 10 March 2021, in final form 8 September 2021

ABSTRACT

Aquaplanet experiments are used to investigate the physical convergence of a Global Storm-Resolving model (GSRM) under successive, two-fold horizontal grid spacing refinements from 160 km to 1.25 km. A methodology based on the Richardson extrapolation method is used with the aquaplanet hemispherical symmetry to quantify convergence. We use the symmetrical and anti-symmetrical solution components to estimate the asymptotic convergence pattern, the asymptotic estimate, and sampling uncertainty. Based on successive refinements, different climate statistics are explored to evaluate if they enter into a convergent regime and, if so, what their convergent value is. Our analysis focuses on global mean statistics related to the general circulation and aspects that influence the climate: the meridional overturning circulation, the tropical structure (the Inter-Tropical Convergence Zone (ITCZ), and the zonal mean thermodynamic state), and the energy and water budget. Our results show a kilometer and hectometer-scale horizontal grid spacing requirement for physical convergence of the meridional overturning circulation structure and global mean statistics. Distinctively, the tropical structure is estimated to be very near their asymptotic values at km-scale grid spacing, but the circulation intensity appears to converge more slowly, as do the storm track and jet-stream. As we increase the horizontal grid spacing, a better representation of clouds and zonal distribution of water vapor drives convergence in the energy and water budget. We conclude that simulations with a resolution of 2.5 km pose a great candidate for multi-decadal simulations within a compromise of the meridional overturning circulation structure convergence and intensity.

PLAIN LANGUAGE SUMMARY

In this study, using the Global Storm Resolving model, ICON, we investigate the effect of increasing horizontal grid spacing on global mean statistics related to the general circulation and aspects that influence the climate: the meridional overturning circulation, the tropical structure, and the energy and water budget. We develop a methodology based on the Richardson extrapolation method to estimate the physical convergence of the model. Our results show an estimated physical convergence at a kilometer and hectometer-scale horizontal grid spacing of the general circulation characterized and global mean statistics. The tropical structure displays physical convergence at kilometer-scale horizontal grid spacing, but the circulation intensity requires finer horizontal

grid spacing, as do the storm track and jet-streams location. Overall, experiments with a horizontal grid spacing of 2.5 km pose as a suitable candidate for multi-decadal simulations within a compromise of general circulation structure and intensity and computation resources required to simulate them.

A.1 INTRODUCTION

We use the Richardson extrapolation method to study the physical convergence of the climate of aquaplanets as we refine the horizontal grid mesh of a global-storm resolving model (GSRM) from 160 to 1.25 km. GSRMs have emerged as a variant to General Circulation models (GCMs), distinguished by their explicit representation of atmospheric moist convection and efforts to minimize the impact of statistical parameterizations (Satoh et al., 2019; Stevens et al., 2019; Hohenegger et al., 2020; Stevens et al., 2020). Through the minimal usage of parameterizations, GSRMs avoid many scale dependencies that complicate physical and numerical convergence studies in traditional GCMs (e.g., Williamson, 2008). Through their explicit treatment of convection, GSRMs offer conceptual simplicity at the price of computational complexity, as they employ kilometer or even hectometer horizontal grid-spacing. Thus, raising the critical question of which horizontal grid-spacing is sufficient and for what.

Physical convergence by horizontal grid spacing of atmospheric characteristics has been investigated in past studies with a special focus on the transient evolution from specified initial conditions. For example, Zarzycki et al. (2019) conducted convergence analyses of the dynamical cores used by large-scale models, including GSRMs, for the case of a splitting supercell test case. They observed a reduction in the sensitivity to resolution at grid spacing of about 500m for bulk properties for each model; however, they also observed a large inter-model spread, suggesting that differences within the model's physics implementation partly determine the convergent solution of the transient dynamics. It is unknown how this difference will be relevant for their climate. For exploring statistical convergence Langhans et al. (2012) used a regional climate model in the European alpine region from 4.4 to 0.55 km horizontal grid spacing for a period of nine simulated days, observing a systematic convergence of bulk properties and the diurnal cycle of precipitation towards 0.55 km horizontal grid spacing. Increasing the domain to global scale simulations, Kajikawa et al. (2016) explored the dependency of deep convection to resolution using the Non-hydrostatic ICosahedral Atmospheric Model (NICAM) (Tomita et al., 2005; Miura et al., 2007b; Sato et al., 2008) for a realistic boundary conditions and a horizontal grid spacing raging from 14 to 0.86 km for 12 simulated hours. The study revealed an apparent convergence at about 1 km horizontal grid spacing for convective cell numbers and no resolution dependency in global statistics. More recently, Hohenegger et al. (2020) used the ICOsahedral Nonhydrostatic (ICON) with the DYAMOND inter-comparison project configuration (Stevens et al., 2019) with horizontal grid spacings varied from 80 to 2.5 km to investigate physical convergence over 40 simulated days. They quantified convergence using the inter-model ensemble spread, observing a convergence towards 2 km horizontal grid spacing for bulk and general circulation characteristics; however, the differences across models participating in the DYAMOND project are still substantial, reflecting that 40 days cannot capture climate statistics robustly, or that the models converge to statistically different climates.

Williamson (2008) investigated convergence using GCMs with the same parameterization suit for horizontal grid spacing raging from 2.8 to 0.35 deg for a period of 14 simulated months. He used aquaplanet experiments, which preserve basic features of Earth's general circulation (Blackburn et al., 2013; Medeiros et al., 2015, 2016), while offering a simplified framework without complex boundary interactions. Aquaplanets have proven themselves as capable of providing valuable insights into phenomena such as tropical circulation and variability (Möbis and Stevens, 2012; Medeiros et al., 2016; Popp et al., 2020; Rios-Berrios et al., 2020), how clouds respond to warming (Landu

et al., 2014; Talib et al., 2018), and the response of other aspects of the climate to external forcing (Stevens and Bony, 2013b; Medeiros et al., 2015; Retsch et al., 2019). Strikingly, Williamson (2008) did not observe convergence with increasing horizontal resolution in bulk and zonal structural characteristics. This non-convergent behavior was attributed to the scale-dependent interplay between small and large scales and parameterization scale dependencies, highlighting how GCMs lack a basic feature of physical simulations that converge to a given answer with improved resolution. Scale aware parameterizations could ameliorate this problem, but also the basic equations would have to be reformulated to avoid the quasi-static assumption (Eliassen, 1948).

In this study, we follow the approach of Williamson (2008) to investigate how different characteristics of the large-scale climate change as the horizontal grid spacing is progressively halved, from 160 km to 1.25 km, but we use a GSRM with the minimum set of parameterizations for radiation, microphysics and turbulence. By avoiding the use of convective parameterization and through the use of a simple turbulence closure, the convergence targets the LES limit. We adopt the aquaplanet configuration for two reasons, both of which were also likely motivations for the earlier study by Williamson (2008). First, they pose a simplified framework that requires a reduced sampling time to reach the mean climate state for convergence assessment; thus, a reduced simulated time. Second, hemispheric asymmetry provides a means to quantify sampling errors due to a under-sampling.

We develop a methodology to evaluate convergence based on the Richardson extrapolation method and hemispherical characteristics of aquaplanet experiments, as detailed in section A.2. Our method helps to formalize convergence and, thereby, to better quantify the behavior of the solutions as horizontal grid spacing is reduced. We emphasize the mean climate characteristics and structural features of the general circulation. For instance, we utilize the Inter-Tropical Convergence Zone (ITCZ) and meridional overturning circulation as characteristics of the general circulation structure and the cold point tropopause in the tropics as characteristics of the tropical vertical structure. Additionally, we explore the effect of grid spacing on cloudiness as measured by the vertical distribution of cloud liquid water and cloud ice in the tropics, and the zonal structure of cloud radiative effect.

Section A.2 introduces the GSRM model configuration, experimental strategy, and our convergence methodology. The physical convergence results of global mean statistics, general circulation, and vertical tropical structure are detailed in sections A.3.1, A.3.2, and A.3.3, respectively. In section A.4 we summarize and discuss our findings.

A.2 METHOD

A.2.1 *Model*

For the GSRM we adopt the ICOsahedral Nonhydrostatic (ICON) model, which was initially developed in a partnership of the Max-Planck Institute of Meteorology (MPI-M) and the German Weather Service (DWD) and is now further developed and maintained in a larger consortium also including the German Climate Computing Centre (DKRZ), the Swiss Federal Institute of Technology (ETH), MeteoSwiss, and the Karlsruhe Institute for Technology (KIT). For the atmospheric component of ICON the non-hydrostatic equations for the atmospheric flow are solved on an triangular grid, which is based on a refined icosahedron, allowing for a quasi-uniform global mesh (Gassmann and Herzog, 2008; Wan et al., 2013; Zängl et al., 2015). ICON employs an Arakawa C-type stenciling

of mass and momentum and its grid structure allows a flexible configuration of the simulation domain with combinations of several earth system components.

A.2.2 *Experiment Setup*

The atmosphere component of ICON (version 2.6.5[1]) is used in a global aquaplanet configuration, retaining only parameterizations for the radiant energy transfer, which is based on RRTMGP (Rapid Radiative Transfer model for General circulation model applications, Parallel, Pincus et al., 2019), a single moment bulk microphysical parameterization consisting of five condensate habits (Baldauf et al., 2011), and 3D turbulent mixing as described in Smagorinsky (1963) with modifications by Lilly (1962) as implemented in the ICON model by Lee et al. (2022a) following an earlier implementation by Dipankar et al. (2015). Further details of the present configuration are provided by Hohenegger et al. (2022). We use a stretched vertical grid with 90 levels, where levels are more finely spaced close to the surface than at the model top at 75 km, where also a damping is employed increasing in strength from 44 km to upwards. In a separate study (Schmidt et al., 2023) explored the effects of vertical resolution and showed that there was surprisingly little sensitivity to the vertical grid spacing for grid spacing finer than what is used in this study. This configurations is employed consistently for all experiments, except for shortening the time-step for higher horizontal grid spacings (Table A.1).

Table A.1: Summary of experiments and model performance using the supercomputer of Levante (DKRZ). For each experiment we present the ICON grid notation (R02Bk), the horizontal grid spacing as the square root of the mean horizontal grid cell area, time step, initialisation strategy, spin-up time, analysis time, the number of node hours required to simulate one day (computational cost), and simulated days per day (simulation throughput).

Grid	Δx	Δt	Initialisation	Spin-Up	Analysis time	Nh/SD	SDPD
	km	s		d	d		
04	160	480	Cold	180	1980	0.015	$6.3 \cdot 10^3$
05	80	360	Cold	180	1980	0.037	$5.1 \cdot 10^3$
06	40	240	Cold	180	900	0.293	$1.3 \cdot 10^3$
07	20	120	Hot	60	480	1.832	405
08	10	60	Hot	60	240	14.72	209
09	5	40	Hot	60	120	84.50	73
10	2.5	20	Hot	60	120	668.4	18.4
11	1.25	10	Hot	50	20	5600	3.9

We perform a series of aquaplanet experiments with horizontal grid spacing ranging from 160 km to 1.25km (Table A.1), whereby this is measured by the parameter Δx defined as the square root of the mean horizontal grid cell area. The experiments performed follow a subsequent refining of the icosahedral grid following the ICON notation **RnBk** (Zängl et al., 2015; Giorgetta et al., 2018), in which each level of refinement increase by four times the cell count in the global horizontal grid. These experiments

[1] icon/icon-2.6.5-rc - hash 4dd46bd54

have been performed on the CPU partition of the supercomputer Levante, which is operated by the German Climate Computing Center (DKRZ).

Our experiments follow the protocol of the AquaPlanet Experiment (Blackburn et al., 2013), and consist of an Earth-sized planet with an Earth-like atmosphere whose lower boundary condition is consistent with a water (no sea-ice) covered surface with a prescribed and zonally and temporally constant sea surface temperature (T_s) (Neale and Hoskins, 2000). In addition, it maintains a perpetual equinox with symmetric-constant irradiation about the equator by setting the eccentricity and obliquity of earth to zero, and using a solar constant of 1361 W/m^2. Ozone follows a constant zonally symmetric profile with respect to the equator, the greenhouse gases are well mixed, and the interaction between aerosol and radiation is neglected.

For T_s all experiments follow the QOBS specification of latitudinal variations following (Neale and Hoskins, 2000).

$$T_s(\phi) = \begin{cases} \frac{27}{2}\left(2 - \sin^2\left(\frac{3\phi}{2}\right)\right)\left(1 + \sin^2\left(\frac{3\phi}{2}\right)\right) & \text{if} \quad |\phi| \leq \frac{\pi}{3} \\ 0 & \text{if} \quad \text{otherwise} \end{cases} \qquad [K] \qquad \text{(A.1)}$$

This idealized approximation of the observed latitudinal variation of sea-surface temperatures most closely resembles Earth's zonally averaged distribution of T_s (Williamson et al., 2013). By imposing a symmetric zonal constant T_s, the resulting aquaplanet is hemispherical symmetric in its mean climate state.

A.2.3 *Accelerating the approach to stationarity*

We define a cold start as a simulation initialized from an idealized temperature sounding of the atmosphere, with the atmosphere at rest and without clouds. The cold start initialization requires a minimum of six to twelve months to overcome the spin-up phase (Blackburn et al., 2013; Webb et al., 2017); only then do we purport to sample the stationary climate, free from the lingering effects of the particular choice of initial data.

Because the computational burden of the spin-up phase considerably increases as we increase the horizontal grid spacing (Table A.1), an alternative form of initialization is chosen for most of the experiments. In these we initialize our numerical experiments by using initial conditions taken from a similar simulation that has already been spun up (Blackburn et al., 2013), albeit on a coarser grid. We refer to this procedure as a hot start. For the hot start, in our experiments, the information flows from coarser to finer horizontal grid spacing (Figure A.1a) as a cascade.

We test the hot start approach by comparing it to a cold start for the 40 km configuration. At this grid-spacing, we observe the transition from a single to a double ITCZ (Figure A.6), and so it is expected to be particularly sensitive to differences in the initialization. We specifically focus on precipitation, as it encapsulates the response of the radiative budget and the general circulation. It is also one of the variables that require the most extended time to reach a state of stationarity (not shown). We identify the end of the spin-up phase by the distance of global mean deviations from the long term mean. Large deviations are taken to be indicative of an ongoing spin-up. Not unexpectedly global means tend to stabilize more swiftly compared to the zonal mean structure. Therefore, we evaluate, in addition to the global means, the stationarity of the

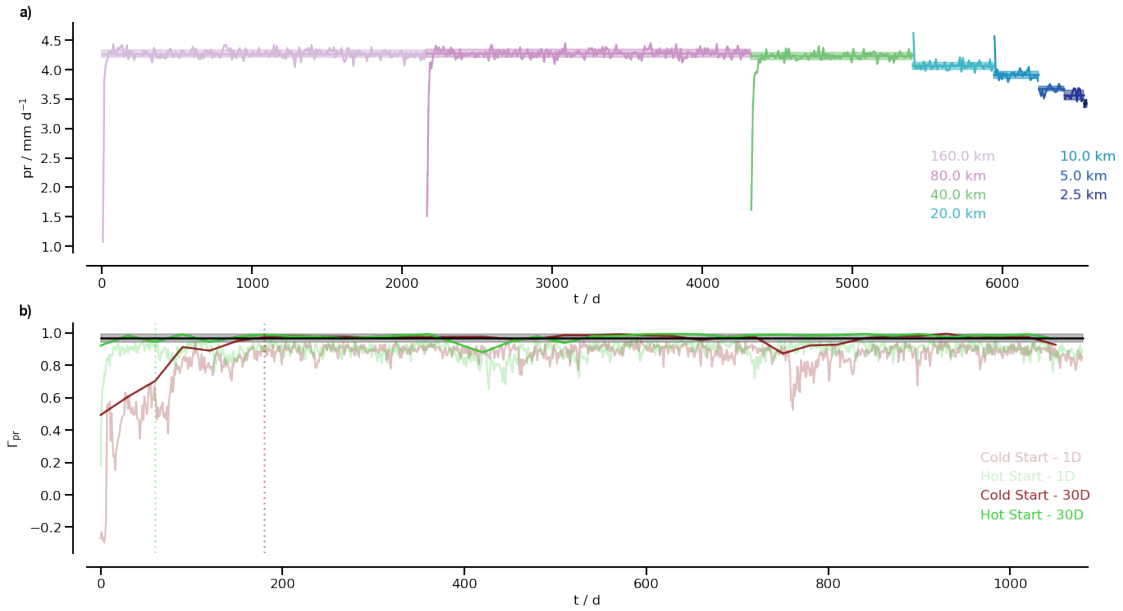


Figure A.1: Spin-up strategy (a) global mean precipitation versus the number of simulated days for each experiment in different line color, the stationary global mean in solid color line within one standard deviation and five-day window means, (b) hot and cold start for the 40 km horizontal grid spacing experiment using Γ_{pr} as a measure of stationarity; in solid lines fifteen-day window means with its respective mean within one standard deviation in black, in shaded lines daily means, and vertical lines indicates 60, 90, and 180 days, respectively.

zonal mean structure as measured by its variance, through the explained variance score defined as follows:

$$\Gamma(y, \hat{y}) = 1 - \frac{Var(y - \hat{y})}{Var(\hat{y})}. \tag{A.2}$$

Here y and $\hat{y}$ are the sample and long term zonal mean hemispherical symmetrical components. Γ gives us a score of how close the reference and a sample are, with a score between zero and one.

Figure A.1b shows the evolution of the zonal structure for both cold and hot start scenarios for precipitation. For the hot start, the spin-up time reduces notably from six to less than two months, and both approaches reach the same stationary state, indicating that no bias was introduced. We also observe a slight and consistent increase in spin-up time as we halve the horizontal grid using both methods. This dependency is not unexpected, as higher horizontal grid spacings resolve smaller scales and increase the variability (not shown). For the 2.5 km horizontal grid spacing experiment, our analysis indicates a spin-up time of about two months. Due to computational limitations, we begin analyzing the results after 50 days for the 1.25 km horizontal grid spacing experiment as a trade off between sampling error and residual biases from the particular choice of the initial state.

A.2.4 *Sampling Time of the Stationarity State*

By design, an aquaplanet experiment with hemispheric symmetric forcing is expected to have a hemispheric symmetric stationary state; hence, as we sample the stationary state, we expect to reach hemispherical symmetry. We take advantage of this assumed property

in two ways. Firstly, we use the hemispherical symmetrical component to augment the sampling statistics. Secondly, we use the hemispherical asymmetrical component to evaluate stationarity for the zonal structure and as an uncertainty band for a system whose stationary state is under sampled.

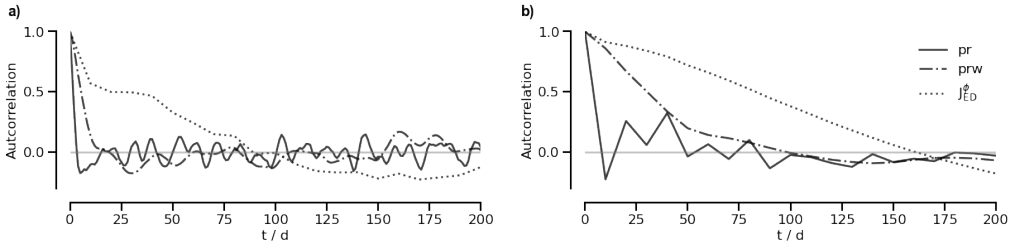


Figure A.2: Sampling analysis of global precipitation (pr), precipitable water (prw) and Eddy-driven jet latitude (a) subsampling time, n estimated as the decorrelation timescale of daily means, (b) sampling time, N, estimated from the cumulative subsample means.

By definition, for a system in stationarity its statistics become independent of the time at which an average is calculated. Medeiros et al. (2016) found that a two-year analysis period measured by sub-monthly sampling provided robust statistics for aquaplanet experiments. However, two years is too expensive for the range of grid-spacings we wish to explore (Table A.2), and for some quantities may not be necessary. Hence we revisit the Medeiros et al. (2016) approach by defining the subsampling time (n) using the autocorrelation of successive daily means and define the sampling interval (N) for the cumulative mean to approach a stationary value. For these purposes we use the 20 km grid-spacing experiment as it exhibits behavior and results similar to higher horizontal grid spacing experiments but with a great many more, 480, analysis days. Once a daily or cumulative subsample mean became de-correlated, we can be confident of the n or N. However, different variables require different n and N, as observed in Figure A.2. Global metrics such as precipitation and precipitable water require n to be less than twenty days. Conversely, large-scale circulation characteristics such as the Eddy-driven jet intensity, require a longer n of 90 days and a N of 180 days.

The autocorrelation helps us verify the sampling statistic of scalar quantities, but becomes more ambiguous for vectors, such as the zonal structure of a given field. We verify the stationarity of the zonal structure by measuring the asymmetrical component of the zonal mean average error between cumulative subsamples. Once the cumulative zonal mean of the average error of the asymmetrical component is smaller than the standard deviation between subsamples of the asymmetrical component of the zonal mean, we can be confident of the stationary state of the sampling time (N_η). Another way to verify the zonal structure's stationary state is using the similarity score (Γ) within a threshold, such as 95% (N_Γ).

Variables with large dispersion and close to zero means, such as meridional and zonal winds at 10m, require a large subsampling and sampling time, i.e., more than 50 and 150 days, respectively. Interestingly, they require less time for the zonal structure sampling requirement. In contrast, variables with considerable means and small dispersion require a shorter subsampling, i.e., less than 20 days, but their sampling times vary according to the nature of the metrics and variability across latitudes. Surface pressure has low variability in the tropics, but is considerable from the midlatitudes toward the poles (not shown), requiring a global sampling time of 130 days, but zonal structure sampling requirements are considerably smaller, 30 days. In contrast, precipitation and precipitable

Table A.2: Summary of sampling statistics for different quantities for the aquaplanet experiment of 20 km horizontal grid spacing, which present similar characteristics as higher horizontal grid spacing experiments. We can observe each variable average (μ), dispersion (σ/μ) in percentage, subsample (n) and sampling time (N) of global means or features, and the zonal mean structure sampling time using the similarity score within 95% threshold (N_Γ) of the hemispherical asymmetrical decay (N_η), rounded up to the nearest value in days, for a given statistic.

Name	Units	μ	σ/μ	n	N	N_Γ	N_η
Surface Pressure	hPa	1006.96	1.03e-4	6	126	22	25
Precipitable Water	kg m^{-2}	20.22	0.70	15	79	1	89
Total Cloud Cover	Fraction	0.86	0.30	10	76	1	111
Precipitation	mm d^{-1}	4.07	1.55	4	10	1	71
Ice path	kg m^{-2}	0.11e-1	7.0e-3	4	150	2	92
Liquid water path	kg m^{-2}	0.17	1.92	8	140	1	117
Net SW at TOA	W m^{-2}	211.93	0.40	8	133	1	117
Surface Wind	m s^{-2}	8.03	1.52	6	153	21	87
10 m Meridional Wind	m s^{-2}	-8.6e-3	371.25	54	162	13	88
10 m Zonal Wind	m s^{-2}	-0.29	38.92	56	169	17	73
ITCZ Intensity	mm d^{-1}	14.79	6.15	11	101	-	-
ITCZ Location	deg	9.53	4.52	14	32	-	-
ITCZ Width	deg	13.24	2.98	20	152	-	-
Storm Track Intensity	mm d^{-1}	4.23	8.18	87	152	-	-
Storm Track Location	deg	43.78	3.71	19	182	-	-

water present significant latitudinal variability, requiring shorter global sampling times compared with its zonal structure sampling requirement, which is larger than 80 days. We use the ITCZ and storm track features (location, width, and intensity) as proxies of the general circulation. The ITCZ requires subsampling time around 20 days and sampling time up to 150 days, while the storm tracks require a longer sampling time of 180 days. For most statistics and characteristics of the general circulation, 120 days is sufficient. For statistical quantities that do not achieve the required sampling time the asymmetrical component provides an estimate of uncertainty.

A.2.5 *Convergence analysis*

A number of prior studies have attempted to quantify the convergence of global models as the grid spacing is refined (e.g., Williamson, 2008; Landu et al., 2014; Retsch et al., 2019; Hohenegger et al., 2020). In these studies, a quantity from a coarse grid solution is compared with solutions from the finest horizontal grid spacing available; this method assumes that the finest grid spacing experiment is closer to the truth. A drawback of this approach is that it cannot ensure that the finest grid spacing experiment is converging toward the truth, and hence closer to it. While these approaches are informative when

quantities converge, they do not provide insight into non-converged quantities, or whether and at what scale they start converging.

The Richardson extrapolation method, widely used in the computational fluid dynamics (CFD) community (Phillips and Roy, 2014), provides an alternative approach, which we adopt here. The Richardson-extrapolation method uses lower-order discrete solutions to derive a higher-order approximation, enabling us to estimate the discretization error:

$$\varepsilon_h = \alpha_p h^p + \text{HOT} \approx \alpha_p h^p \tag{A.3}$$

where HOT describe high-order terms, h is the grid spacing, α_p a constant, and p the formal order of convergence, which depends on the discretization scheme used. Assuming that the numerical solution is in the asymptotic range of convergence, we can drop the HOT and approximate the discretization error of two systematically refined grids with asymptotic numerical solutions:

$$\varepsilon_h = \frac{u_{rh} - u_h}{r^p - 1} \tag{A.4}$$

where u is a characteristic or metric to evaluate between the two simulations with grid spacing h and rh respectively, with r the grid refinement factor and h is, therefore, the coarsest grid spacing between the two evaluated experiments.

Because we do not know the formal order of accuracy, we estimate the order of accuracy (p) using an additional refined grid (r^2h):

$$p = \min\left(0.5, \frac{\ln\left(\frac{u_{r^2h} - u_{rh}}{u_{rh} - u_h}\right)}{\ln(r)}\right) \tag{A.5}$$

where 0.5 is a limiting factor to minimize unrealistic large error estimates (Phillips and Roy, 2014) that may emerge in non-convergent regimes. Once we estimate the order of accuracy, we estimate the asymptotical value by:

$$\tilde{u} = u_{r^2h} + \frac{u_{r^2h} - u_{rh}}{r^p - 1} \tag{A.6}$$

Hence our application of the Richardson extrapolation method requires three experiments with grid spacing [h, rh, and r^2h] or horizontal grid spacing of [Δx, $0.5\Delta x$, $0.25\Delta x$], where r is 0.5 since we halve the grid as we refine the horizontal grid spacing. The Richardson extrapolation method is valid if the three evaluated metrics are in their asymptotical regime, meaning that Eq. A.4 is equal to the formal order of accuracy. Since we calculate the order of accuracy, we adopt a comparative approach, which assesses the discretization error growth between two consecutive pairs (ε_h and ε_{rh}) and whether it follows the discretization convergence rate Eq. A.3. In the asymptotic regime α_p should be constant. By evaluating $\alpha_h = \varepsilon_h / h^p$ for each pair (ε_h and ε_{rh}) and forming their ratio, we create the metric:

$$\zeta = \frac{\varepsilon_h(p)}{\varepsilon_{rh}(p)} r^p \tag{A.7}$$

where $\zeta \approx 1$ indicates that we are in the asymptotical regime.

Since we estimate p from the output, which is noisy, we use two alternative ways to measure confidence in our results and facilitate convergence analysis. First, we define a range of confidence based on ζ, where $|1 - \zeta| < 0.1$ is considered highly reliable,

$\zeta < 0.25$ moderately reliable, $\zeta < 0.5$ questionable, and $\zeta > 0.5$ unreliable. Second, we assess uncertainties arising from undersampling by evaluating each metric within the hemispherical symmetrical component $\hat{u}$ and its uncertainty by re-calculating it within a plus and minus hemispherical asymmetrical component δ_u, obtaining an uncertain range $u = \hat{u} \pm \delta_u$. We use the functional approach (Hughes and Hase, 2012), corroborated with the Monte Carlo approach (Possolo and Iyer, 2017), to propagate the uncertainty δ_u through Eq. (A.6) and obtain the estimated asymptotic value ($\tilde{u}$) within its respective uncertainty range ($\delta_{\tilde{u}}$). We perform this analysis for every three sets of contiguous horizontal resolution experiments and evaluate the evolution of the $\tilde{u}$ and p with their respective confidence and uncertainty bands.

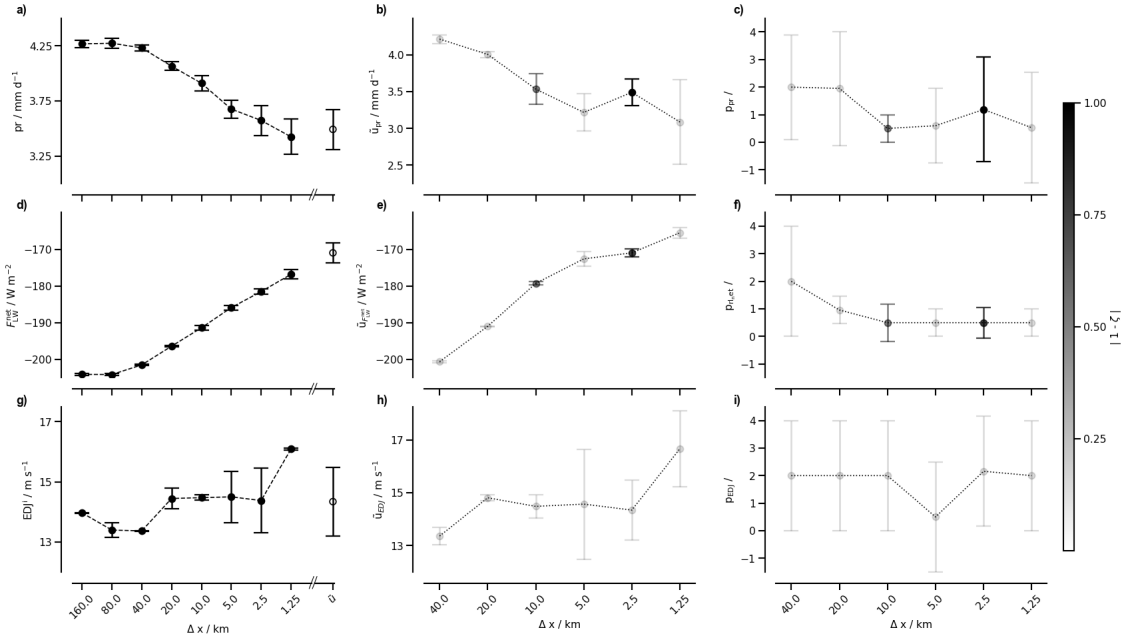


Figure A.3: Convergence analysis of three distinctive convergence patterns, one for each row: reliable convergence of global mean precipitation (a-c), steady convergence pattern of global mean atmospheric net longwave radiation (d-f), and unclear convergence pattern of Eddie-driven jet intensity (g-i); with three different characterizations within their respective uncertainty bars and confidence (ζ): metric at each horizontal grid spacing and respective asymptotical estimate ($\tilde{u}$) at 2.5 km horizontal resolution (a, d, g), asymptotical estimate ($\tilde{u}$) within three consecutive horizontal grid spacing (b, e, h), and the estimated order of accuracy (p) within three consecutive horizontal grid spacing (c, f, i).

Figure A.3 shows the convergence analysis for three distinctive convergence classes: reliable, steady, and unclear. A reliable convergence pattern (Figure A.3a,b,c) displays a steady convergence towards $\tilde{u}$ as we increase the horizontal grid spacing, reaching the asymptotical estimate at 2.5 km horizontal grid spacing, While a steady convergence pattern (Figure A.3d,e,f) displays a similar pattern but has yet to reach $\tilde{u}$ at 2.5 km horizontal grid spacing. In both cases, the asymptotical estimates across resolutions show slight differences between each estimate, and they lie between their respective uncertainty bars as we increase the horizontal grid spacing with moderate or greater confidence ($|1 - \zeta| < 0.25$). In contrast, an unclear convergence pattern (Figure A.3g,h,i) might display a visual convergence or a regime change as we increase the horizontal grid spacing, making it vague if the metric has arrived at convergence. Its $\tilde{u}$ across resolutions show low confidence ($|1 - \zeta| > 0.25$) and large uncertainty bars within resolutions;

hence, $\tilde{u}$ might lie in the range sampled, but we cannot be confident. Additionally, we evaluate p, the rate of convergence of each metric, which indicates how fast the metric converges as we increase the horizontal resolution. For highly confident variables, we can estimate the required resolution to achieve a desired error, i.e., standard deviation within ten-day sampling time (Table A.2), using Eq. A.3; however, it must be considered as an estimate.

A.3 RESULTS

We evaluate convergence with and without the 1.25 km horizontal grid spacing given the limited (20-day) sampling time of the finest grid simulations. We assess convergence confidence (reliability) assessed using ζ. Results are described as converged when the quantity in question systematically reaches its asymptotical estimate within its uncertainty bars, and the 1.25 km is within the asymptotical uncertainty bars for those cases where a 20-day sampling time is sufficient to sample the quantity (Table A.2). Metrics with a systematic convergence pattern display convergence but have yet to achieve asymptotical estimates within their uncertainty bars. Metrics might display oscillatory convergence, for which we verify if the uncertainty bars are within the asymptotical estimate, suggesting potential convergence with an extended sampling. We consider estimates of the quantity as diverging if the deviations grow as we refine the horizontal grid spacing. Lastly, we consider unclear convergence patterns in the case of large uncertainty bars and low confidence ($|1 - \zeta| > 0.25$) as we increase horizontal grid spacing. In this fashion the 1.25 km horizontal grid spacing serves as a further check on convergence estimates.

Below, we first explore convergence for global mean statistics (§3.1), followed by metrics that characterize main features of the general circulation (§3.2), and, lastly, §3.3 metrics that characterize the vertical structure within the tropics.

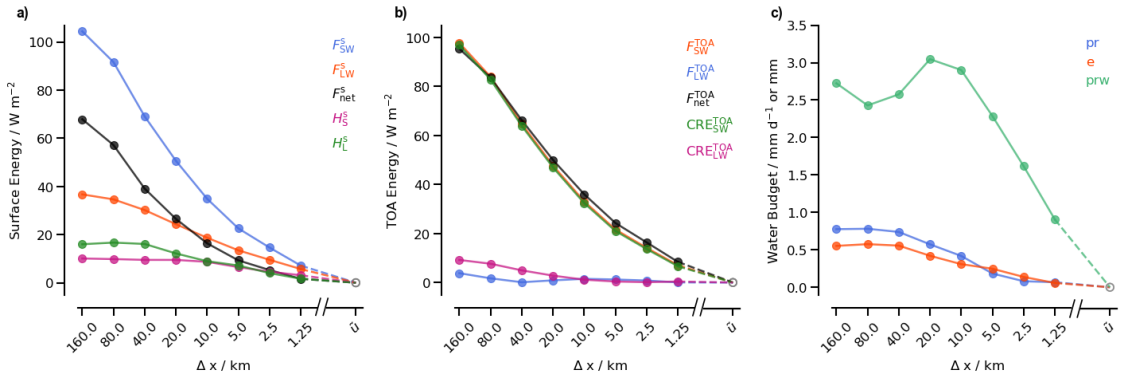


Figure A.4: Global mean convergence analysis expressed as the absolute relative difference to asymptotical estiamte ($\tilde{u}$) across resolutions (x-axis): (a) surface components of the energy budget, (b) TOA component of the energy budget and cloud radiative effect, (e) water budget composed by precipitation (pr) and evaporation, and precipitable water(prw)

A.3.1 *Global Mean Statistics*

Figure A.4 shows energy and water budget components for different horizontal grid spacing relative to their respective asympotical estimate ($\tilde{u}$). We observe different

convergence rates for different quantities; the precipitable water path (prw) exhibits a change of convergence regime at 10 km horizontal grid spacing, and the shortwave component of the energy budget displays a much higher sensitivity to horizontal grid spacing compared to the longwave component, a characteristic also observed by (Hohenegger et al., 2020).

We define the energy fluxes with a positive sign as atmospheric heating, i.e., the net longwave radiation at TOA and net shortwave radiation at the surface are negative. The longwave cloud radiative effect at the top of the atmosphere ($\mathrm{CRE}_{\mathrm{LW}}^{\mathrm{TOA}}$), precipitation (pr), evaporation (e) or latent heat fluxes ($H_{\mathrm{L}}^{\mathrm{s}}$), and vertically integrated cloud ice (clivi) show a reliable convergence at a 2.5 km horizontal grid with an estimated order of convergence greater than 1. The net shortwave and longwave radiation at the surface ($F_{\mathrm{SW}}^{\mathrm{s}}$ and $F_{\mathrm{LW}}^{\mathrm{s}}$), shortwave cloud radiative effect ($\mathrm{CRE}_{\mathrm{SW}}^{\mathrm{TOA}}$), net shortwave radiation at the top of the atmosphere ($F_{\mathrm{SW}}^{\mathrm{TOA}}$), sensible heat flux ($H_{\mathrm{S}}^{\mathrm{s}}$), precipitable water (prw), and vertically integrated cloud liquid water (cllvi) display a steady high-reliability convergence pattern but have yet to converge at a 1.25 km horizontal grid spacing with a sublinear ($p \approx 0.6$) order of convergence, requiring hectometer horizontal grid spacing for convergence. In contrast, net longwave radiation at the top of the atmosphere ($F_{\mathrm{LW}}^{\mathrm{TOA}}$) displays an unclear, low-reliability, convergence pattern with an apparent convergence at 2.5 km horizontal grid spacing, but at 1.25 km, deviations (not shown) affirm our lack of confidence in the estimate.

Two key factors influence the convergence rate of the global statistics. First are clouds, wherein a consistent reduction of shallow clouds and an increase in clear-sky conditions as we increase the horizontal grid spacing is observed. Similar behavior has been noted in earlier studies by Kajikawa et al. (2016) and by Hohenegger et al. (2020) at global scales, and by Stevens et al. (2020) for a large domain limited area simulation over tropical oceans. The reduction of shallow clouds at higher resolutions contributes to a notable increase in F_{SW}^{s} and F_{SW}^{TOA}. A clear convergence tendency is evident, but appears to require horizontal grid spacings on the order of a hectometer; consistent with evidence from large-eddy simulations, (e.g. Stevens et al., 2020) evaluated against observations. The increase in clear-sky conditions with refinement also contributes to a decrease in cloud radiative heating, reducing the longwave downwelling and increasing the net upwelling at the surface. At first glance, the F_{LW}^{TOA} seems to converge as we refine the horizontal grid spacing. Upon close inspection, F_{LW}^{TOA} reaches its maximum at a horizontal grid spacing of 10 km, consistent with decreased precipitable water up to 20 km horizontal grid spacing and the circulation shift. One would expect to observe the maxima of net longwave radiation in unison with the precipitable water minimum, but the circulation shift compensates for its effect (Figure S1 and S2). Together, they measure shifts in the circulation – one being the transition from a single to a double ITCZ at 40 km horizontal grid spacing and a progressive poleward shift of the storm tracks – to be discussed further below.

A.3.2 *The Meridional Overturning Circulation*

Figure A.5 presents the zonal mean mass meridional stream function and the zonal wind for two horizontal grid spacings (40 km and 2.5 km). It highlights the structure of the mean overturning circulation, with easterlies in the tropics and westerlies in the extra-tropics, along with their respective jets. The Hadley cell, delimited by where the stream function and near surface winds change sign is readily evident.

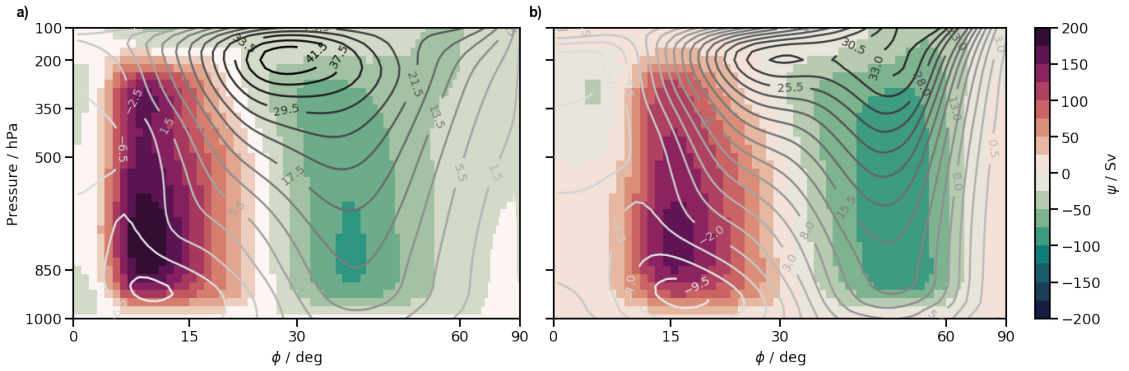


Figure A.5: Vertical structure in pressure levels of the time-averaged zonal meridional stream function and zonal mean winds as contours with intervals of 2.5 ms^{-1} for the 2.5 km resolution for two horizontal grid spacings: (left) 40 km and (right) 2.5 km.

As the grid is refined the circulation weakens and moves poleward. The transition happens around a grid spacing of 40 km, as is evident (Fig. A.6) in both the surface precipitation (pr) and the zonal mean meridional mass stream function (ψ) at 500 hPa. At grid spacings finer than 20 km, the zonal structure appears relatively stable to further refinement of the horizontal grid. These results differ from earlier studies using NICAM (Tomita et al., 2005), which featured a single ITCZ at all resolutions. One difference with the earlier work is that they used the CNTL rather than QOBS specification for the SST, which would favor a single ITCZ. Nonetheless preliminary tests suggest that ICON maintains a double ITCZ at fine grid-spacings, even when using the CNTL SST (not shown). Other factors, such as cloud radiative effects, are apparently also important.

To be quantitative, we characterize the ITCZ zonal structure by its location (superscript ϕ), width (superscript w), and intensity (superscript i) as measured either in terms of precipitation (subscript pr), or stream function (subscript ψ). With this notation, $\mathrm{ITCZ}^{\mathrm{w}}_{\psi}$ denotes the width of the ITCZ as determined by the crossing zero of the stream function and $\mathrm{ITCZ}^{\mathrm{w}}_{\psi}$ by the crossing zero of precipitation minus evaporation. Based on these measures, the ITCZ location and intensity converge around a 2.5 km horizontal grid spacing, but the convergence of the ITCZ width, as measured either by the stream function ($\mathrm{ITCZ}^{\mathrm{w}}_{\psi}$) or the precipitation ($\mathrm{ITCZ}^{\mathrm{w}}_{\mathrm{pr}}$), is more ambiguous. In the case of $\mathrm{ITCZ}^{\mathrm{w}}_{\mathrm{pr}}$, the reliability of convergence is questionable, and in the case of $\mathrm{ITCZ}^{\mathrm{w}}_{\psi}$, a low rate of convergence ($p \approx 0.5$) and a lack of convergence at 1.25 km indicates that hectometer horizontal grid spacing may be necessary for convergence. We have also evaluated the equatorial atmospheric energy input (AEI_0) for each of the resolutions (Figure A.7). Its changes are consistent with changes in the position of the ITCZ, as it changes sign as the ITCZ shifts from a single ITCZ which exports energy from the convective regions.

The shift from a single ITCZ to a double ITCZ at 40 km horizontal grid spacing also can be seen in terms of the diagnostic atmospheric energy framework, which associates the sign of the equatorial atmospheric energy input with a single or double ITCZ (e.g., Adam et al., 2016; Bischoff and Schneider, 2016; Adam, 2021). As the ITCZ shifts from a single to a double ITCZ this input shifts from a positive to negative value, highlighting that the circulation transports energy out of the convecting regions. In contrast to (Adam et al., 2016), the ocean response plays no role in our simulations with fixed SSTs. Moreover, given that the boundary conditions for all the simulations are the same, it shows that the atmospheric energy diagnostic is more a consequence than a cause of where the convection locates. Atmospheric cloud radiative effects associated with the

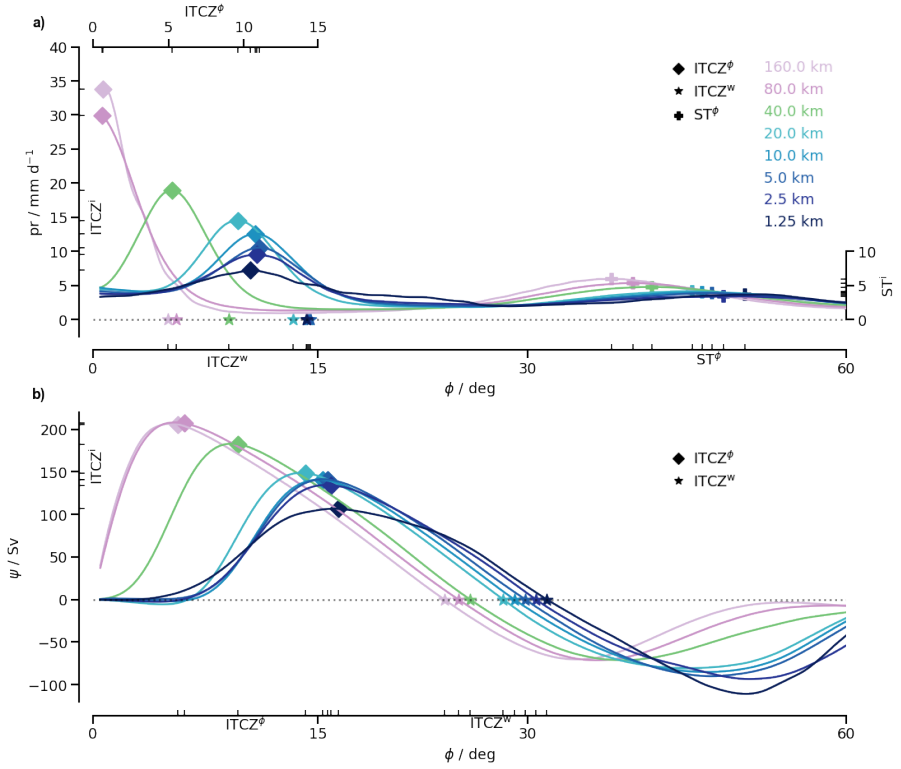


Figure A.6: ITCZ zonal profile across resolutions (a): precipitation (pr), and (b): zonal-mean mass meridional stream function (ψ). The ITCZ, defined by its latitude and width, are marked with a cross and star, respectively. The storm tracks latitude with a diamond for each horizontal grid spacing. On the x, y, and extra axes, we identify the ITCZ latitude (ITCZ$^{\phi}$), width (ITCZ^{w}), intensity(ITCZ^{i}), and storm tracks latitude (ST$^{\phi}$) and intensity (ST^{i}) for each horizontal grid spacing.

ITCZ (Figs.S3, S4) are large (35 Wm^{-2}) and the changes in the net energy input at the equator as the convection moves off the equator, can largely be explained by changes in the long-wave radiation at the top of the atmosphere (Figs.S3, S4). What was surprising in our simulations is the convergence of equatorial atmospheric energy input to near zero (already at horizontal grid spacings of 10 km), indicating that subsidence can balance the radiant heat loss in the interior, and surface fluxes balance the loss of radiant energy in the boundary layer, with the circulation neither importing nor exporting energy from the extra-tropics.

Accompanying the increasingly poleward extension of the Hadley's cell edge, we observe a poleward shift of the subtropical jetstream, the maximum surface pressure, and the change in the direction of the surface winds from easterly to westerly latitude. The subtropical jetstream, identified at the maximum vertical shear of the horizontal winds between 200 hPa and 850 hPa (Adam et al., 2018), weakens and moves polewards with grid refinement (Figure A.5). Its intensity and latitude display an unclear convergence with apparent visual convergence at 5 km horizontal grid spacing but with unclear confidence and large uncertainty bars in their asymptotic estimates (not shown) and an estimated rate of convergence of 2. The maximum surface pressure and the change in the direction of the surface winds from easterly to westerly latitude, which are proxies of the boundaries of the tropics (Adam et al., 2018), also exhibit an unclear steady convergence pattern, with results for the 1.25 km not yet convergent, and a low estimated rate of convergence of 0.5.

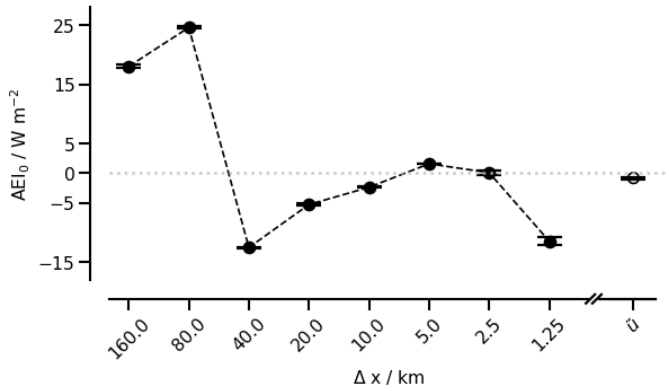


Figure A.7: Atmospheric net energy intake at the equator (AEI_0) across resolutions

To quantify changes in the extra-tropical circulation we measure the storm-track (ST) position and intensity in terms of the secondary precipitation maximum in the extra-tropics and we measure the eddy-driven jet in terms of the position through the 850 hPa zonal wind maximum. As resolution is refined, storm tracks move poleward and become less intense (Figure A.6), however the great variability of these systems makes the statistical convergence more difficult to quantify. The eddy-driven jet behaves similarly, albeit with similar difficulties in unambiguously establishing convergence. More reliable statements about their convergence with grid refinement will require longer simulations.

A.3.3 *The Zonal Mean Thermodynamic State in the Tropics*

Figure A.8 shows the mean vertical structure (0-30 deg) of temperature (T), relative humidity (RH), cloud liquid water (clw), and cloud ice (cli) for the different experiments.

The vertical thermal structure of the tropical atmosphere converges as we increase the horizontal grid spacing, qualitatively converging at horizontal grid spacings finer than 20 km (Figure A.8a). As it converges it becomes less stable, and the cold-point troposphere warms quite markedly (by nearly 10 K), and descends slightly if at all, as in the latter case the small changes are difficult to separate from variability.

Relative humidity shows a gradual and slow convergence with resolution, with progressive drying of the boundary layer and moistening of the free troposphere (Figure A.8b). More quantitatively, the relative humidity maximum and height delineate the boundary layer depth, both appear to converge at 2.5 km horizontal grid spacing, although more questionably so for the boundary layer height whose small changes are difficult to resolve and separate from variability. The secondary peak in relative humidity in the upper troposphere is still noticeably increasing at km-scale grid-spacings with an estimated convergence rate of 1.4 and 1.3, respectively.

The cloud liquid water maximum and height display a highly reliable convergence pattern but have yet to achieve convergence at 1.25 km horizontal grid spacing with an estimated convergence rate of 1.4 and 0.8, respectively (Figure A.8c). The strong resolution dependence of marine boundary clouds has been noted in other studies, (e.g., Hohenegger et al., 2020; Stevens et al., 2020) and largely explains the grid sensitivity of the planetary albedo. Through a reduction in the downward longwave radiation at the surface (Figure S3), a reduction of marine boundary layer clouds will reduce the radiative cooling of the atmosphere, and would explain the gradual reduction in precipitation

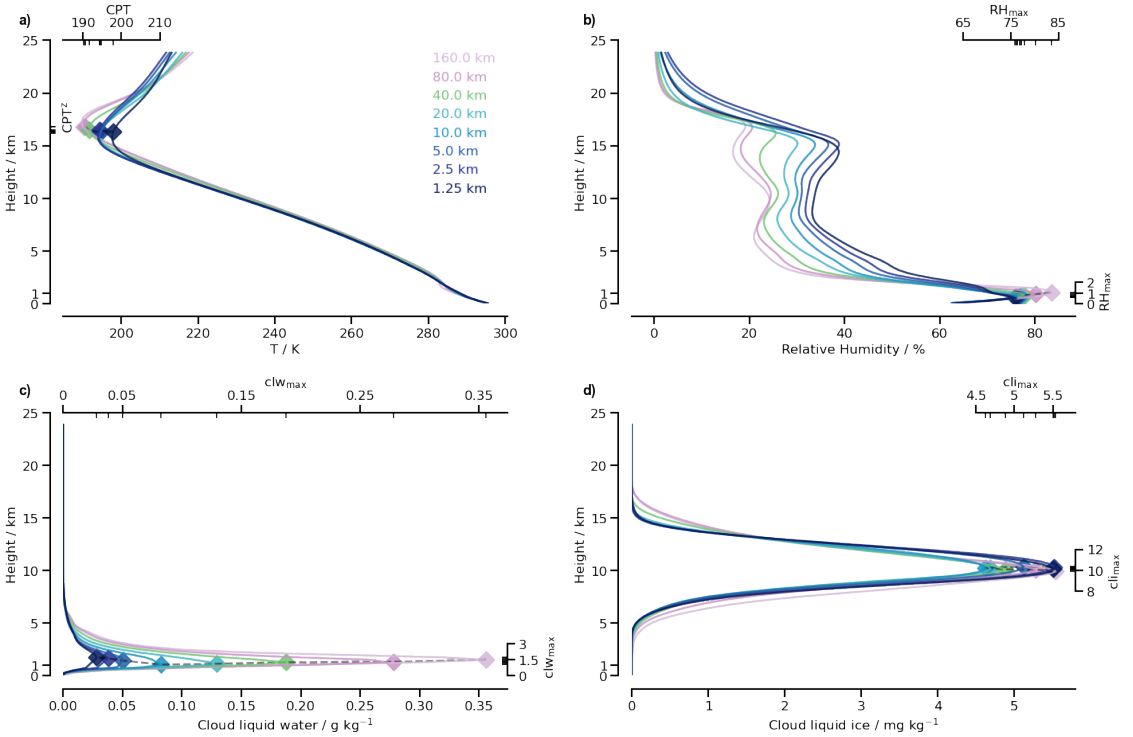


Figure A.8: Tropical (30N to 30S) vertical structure profiles with their respective characterization (metric) (a) temperature characterized and the cold point tropopause (CPT), (b) relative humidity and maximum above boundary layer (RH_{max}), (c) cloud liquid water content and its maximum (clw_{max}), and (d) cloud ice content and its maximum (cli_{max}). On the x, y, and extra axes, we identify each characterization (metric) magnitude and height for each horizontal grid spacing

with resolution. If one combines this cloud desiccation tendency with past work at much finer resolution, it suggests that the planet might darken as trade-wind clouds become more diminutive with finer resolution and converge at hectometer scales (Stevens et al., 2001; Schulz and Stevens, 2023), but as resolution is even refined further, to decameter scales, thin stratiform clouds associated with the detrainment from trade-wind clouds will become better resolved and more pronounced, causing the planet to brighten. This would suggest that the planetary albedo may not be monotone in grid-spacing (or effective viscosity).

Ice clouds also show a lack of monotonicity in their convergence, albeit with much smaller changes, at scales of tens of kilometers (Figure A.8d). Part of this may be related to the changes in the structure of the ITCZ. As the convection transitions from a single ITCZ on the equator to two ITCZs off the equator, ice-cloud amount is reduced, but with progressive refinement in the double ITCZ regime, ice-cloud amount begins to increase again. Convergence in this regime is reliable, but slow, as it has yet to achieve convergence at 1.25 km horizontal grid spacing ($p \approx 0.5$). The height of the maximum in ice-cloud amount is robust to resolution changes, and occurs well below the secondary peak in relative humidity. Given the much larger changes in circulation (Figure A.6) and humidity (Figure A.8b) this suggests that the maximum is mostly controlled by microphysical and thermodynamic processes that are insensitive to resolution.

A.4 SUMMARY AND DISCUSSION

The convergence and resolution effect of the climate characteristics in numerical models is generally based on comparing coarse horizontal grid spacing with the finest horizontal grid spacing achievable or within inter-model spreads (e.g., Williamson, 2008; Landu et al., 2014; Retsch et al., 2019; Hohenegger et al., 2020). However, this methodology cannot address whether horizontal grid spacing is sufficient or how far we are from convergence within a given tolerance. We develop a methodology based on the Richardson-Extrapolation method that sheds light on convergence based on the asymptotic estimate and convergence rate within a confidence interval for the energy and water budget, the meridional overturning circulation, and the thermodynamic zonal vertical structure in the tropics. Convergence is assessed using aqua planet experiments, whose symmetric structure reduces sources of large-scale noise, and their effects on the sampling time required to establish the statistics of the stationary climate, while preserving many of its main features (Blackburn et al., 2013; Medeiros et al., 2015, 2016). We use the asymmetrical hemispherical component as a metric to quantify uncertainties arising from short sampling times and evaluate convergence accordingly. By using the Richardson-Extrapolation method and the asymmetrical hemispherical component as an uncertainty metric, we evaluate the convergence of mean climate characteristics as they asymptotically approach convergence.

As the horizontal grid spacing is progressively refined a single ITCZ transitions toward a double ITCZ, first at horizontal grid spacings of about 40 km. The general features of the zonally averaged circulation show a poleward drift, both in terms of indicators of the tropical rain bands and of the extra-tropical storms, that tends to converge at horizontal grid spacings of 1 km-5 km. Convergence of features in the extra-tropics are, however, more difficult to diagnose due to their more pronounced variability over larger timescales, which we are unable to sufficiently sample. That one might be tempted to argue as to whether one or the other feature converges at resolutions closer to 10 km or closer to 1 km, this merely highlights how far even very large-scale dynamical features are from being resolved in the present generation of climate models, as even the highest-resolution variants of these employ grid spacings closer to 50 km.

Most of the terms of the atmospheric energy budget converge at a horizontal grid spacing of around 2.5 km. However downwelling long wave radiation at the surface and reflected shortwave radiation at the top of the atmosphere still change markedly (by several $\mathrm{Wm}{-}2$ for a doubling of grid spacings) with refinement at these scales. Changes in downwelling longwave radiation are less pronounced than changes in outgoing shortwave radiation. However, these likely explain the progressive reduction in precipitation with finer resolution, as the distribution of precipitable water is nearly converged already at much coarser grid spacings (closer to 10 km). Both the changes in long and shortwave radiation can be attributed to a large sensitivity of liquid water clouds to resolution, something that has also been noted in previous studies (Landu et al., 2014; Kajikawa et al., 2016; Hohenegger et al., 2020; Stevens et al., 2020).

Except for the noted changes in low-cloud amount, the thermal structure of the tropics displays converges at kilometer-scale horizontal grid spacing. This is evident both in terms of its meridional structure (ITCZ) as well as its vertical structure (Figure A.8). The cold point tropopause temperature shows pronounced (10 K) changes with resolution for horizontal grid spacings larger than 20 km but its height and value change relatively little as the grid spacing is further refined. The humidity structure shows a progressive moistening of the free, and particularly upper, troposphere, while the lower troposphere

dries and the boundary layer shallows with progressive refinement of the horizontal grid. Changes in the humidity structure of the upper troposphere are still on the order of a few percent RH as the grid is refined from 5 km to 2.5 km, but changes in humidity structure of the boundary layer appear to be little changed by further refinement at these resolutions. The maximum cloud ice shows some sensitivity to the structure of the ITCZ, but overall varies remarkably little with resolution.

Some caveats of our methodology merit mention. One is the high sensitivity to quantities near zero, or that change very little, with grid spacing. In this situation the reliability (as measured by ζ) in our estimates of the horizontal grid spacing required for convergence within a certain tolerance is low, and estimates of the rate of convergence are unstable. However in this case, where the answers are anyway not changing a lot, we are not particularly concerned by an inability to quantify precise convergence. A second caveat is that when convergence is established, this by no means ensures that we have converged to the correct answer, as the convergent answer may depend on non-fluid dynamical aspects of the simulation in ways that are not representative of reality. Even whether or not quantities converge may depend on the formulation of sub-grid processes, and merits further investigation. Notwithstanding these caveats, this study suggests that if ways can be found to tame the resolution dependence of shallow marine boundary layer clouds, a km-scale climate model can resolve major circulation features, but that for simulations at grid spacings coarser than 10 km methods would also need to be devised to tame a great many other climate relevant features of the general circulation.

OPEN RESEARCH SECTION

The aquaplanet experiments were done with the ICON release candidate 2.6.5 as commit 4dd46bd54. The source code is available at https://doi.org/10.17617/3.1XTSR6, and the ICON model is available to individuals under licenses https://mpimet.mpg.de/en/science/modeling-with-icon/code-availability. The scripts used to plot the figures and process and analyze the data in the paper can be found as supplementary material in the repository: https://doi.org/10.17617/3.NUYNV9.

ACKNOWLEDGMENTS

The authors would like to thank C. Hohenegger, H. Segura, and J. Bao for valuable discussion and feedback We are grateful for advice and technical support from colleagues at DKRZ and MPI-M, specifically from R. Redler and JF. Engels. The research was supported by public funding to the Max Planck Society, the European Union's Horizon 2020 research and innovation program project NextGEMS funded C. Hohenegger and B. Stevens under the grant agreement number 101003470, and the WarmWorld funded by the German Federal Ministry of Education and Research under the funding code 01LK2202B.

B

APPENDIX B

The work in this appendix is to be submitted to the *Geophysical Research Letters* as

Peinado-Bravo, A., Klocke, D., & Stevens, B. “Global Warming and the Robust Response with increasing horizontal grid spacing"

Global Warming and the Robust Response with increasing horizontal grid spacing

Angel Peinado-Bravo[1,2], Daniel Klocke[1],Bjorn Stevens[1]

[1] Max Planck Institute for Meteorology, Hamburg, Germany
[2] International Max Planck Research School on Earth System Modelling, Max Planck Institute for Meteorology, Hamburg, Germany

Manuscript submitted 29 September 2022

ABSTRACT

Previous work has displayed the physical convergence of a Global Storm-Resolving model (GSRM) without convection parameterizations using a new methodology based on the Richardson extrapolation method to assess convergence. The large-scale structure, i.e., the Intertropical Convergence Zone (ITCZ), and global mean statistics require kilometer and hectometer-scale horizontal grid spacing as a better representation of clouds and zonal distribution of water vapor drives convergence in the energy and water budget is observed. What has yet to be clarified is if a GSRM displays physical convergence, does it display a robust response of mean climate statistics in a warming scenario with increasing horizontal grid resolution? At a 10 km horizontal grid spacing, the intensification of the hydrological cycle and longwave climate feedback agrees with other GSRM studies and observations and become insensitive to increasing resolution. In contrast, general circulation response and shortwave climate feedback require higher horizontal grid spacing.

B.1 INTRODUCTION

The increasing available computational power and advances in high-performance computing allow for the usage of kilometer-scale global storm-resolving models (GSRMs) in a wide range of horizontal grid spacing (Satoh et al., 2017; Stevens et al., 2020; Hohenegger et al., 2022). GSRMs have arisen as a promised tool to understand underlying and resolve small-scale processes (Palmer and Stevens, 2019), which are parameterized (subgrid parameterizations) in traditional Atmospheric General Circulation Models (AGCMs). Subgrid parameterization has been attributed to uncertainty in climate response under a warming scenario with relevant inter-model differences (Sherwood et al., 2020; Zelinka et al., 2020, 2022), even in idealized simulations (Stevens and Bony, 2013a). However, increasing horizontal grid spacing by explicitly resolving convection might reduce uncertainty. i.e., in clouds and water-vapor-lapse rate, has yet to be understood. In this way, we aim to comprehend to which extent the small-scale processes aggregate and feedback on global-scale climate features that control the climate response under forcing.

Feedback within circulation, clouds, and local feedback responses have been identified as significant sources of uncertainty in climate change prediction (Voigt and Shaw, 2015). Different climate models present different patterns in the circulation response with food print in the changes of clouds in the tropics and extratropics. To our knowledge, Kodama et al. (2015) and Noda et al. (2019) was the first to perform a long-time integrated (20 years) AMIP (control, +4K, 4xCo2) simulation using NICAM at 14 km without convection parameterization, hoping to represent clouds better. Their finding highlighted a neutral response to the shortwave climate feedback with reduced shallow clouds over the ocean and increased anvil clouds in the deep tropics, opposing the expected iris effect hypothesis (Bony et al., 2016). This pattern has also been observed in a simpler experiment setup, the radio convective equilibrium, which emulates the tropics. However, different models with or without parameterization display different responses, and only a third of the models do not observe the iris effect hypothesis (Silvers et al., 2023). More recently, from the global mean thermodynamic and hydrological response, Merlis et al. (2024b) performed a high-resolution (3 km horizontal grid spacing) AMIP experiment (control, +4K, 4xCo2) using X-Shield for 2 years without convection parameterization and compared with the inter-model spread of CMIP6. Similar to Noda et al. (2019), they found a muted effect in shortwave climate feedback, and in comparison to CMIP6, longwave, earth climate sensitivity, and the hydrological cycle are within the inter-model spread of CMIP6. However, many could argue that 2 years is too short to observe the climate response, and we require longer integration times.

Even though available computational power has increased, the computational cost of investigating climate at kilometer-scale horizontal grid spacing is still significant (Peinado-Bravo et al., 2024). One way to reduce the computational constraint is to use simplified experiments that capture the response of more complex experiments. Medeiros et al. (2015) showed that the response of aquaplanet experiments is similar or equal to that of AMIP-type simulations, thus highlighting its usage in understanding the climate response within shorter integration times. These characteristics have been exploited to investigate, i.e., the interplay within of clouds, water vapor, and circulation (Voigt and Shaw, 2015; Shaw and Voigt, 2016), and cloud radiative feedback (Narenpitak et al., 2017; Retsch et al., 2019), bringing closer the gap to what drives uncertainty within models.

In this research, we utilized aquaplanet experiments to explore the response of various characteristics of the large-scale climate under a forcing scenario as the horizontal grid spacing was progressively reduced from 160 km to 1.25 km. Our approach involved

using a GSRM with a minimal set of parameterizations for radiation, microphysics, and turbulence, while avoiding convective parameterization. By employing a simple turbulence closure, the convergence targets the LES limit. The adoption of the aquaplanet configuration was motivated by several reasons, including the simplified framework requiring reduced sampling time for convergence assessment and the ability to quantify sampling errors due to under-sampling through hemispheric asymmetry. Additionally, the response of aquaplanets is akin to more complex experiments such as AMIP simulations (Medeiros et al., 2015)

The climate response to external forces is a complex and extended topic. To simplify the manuscript, we focus on the same topics and unavoidably cannot cover all aspects. We aim to provide a global overview rather than to focus in specific topics deeply. Thus, we investigate the global mean thermodynamic and hydrological and the general and tropical circulation response. For instance, we utilize the Inter-Tropical Convergence Zone (ITCZ) and meridional overturning circulation as proxies of the general circulation structure and the cold point tropopause in the tropics as characteristics of the tropical vertical structure. Additionally, we explore the effect of grid spacing on cloud response in the tropics measured by the vertical distribution of cloud liquid water and cloud ice, as well as the zonal structure of cloud radiative effect.

B.2 RESULTS

Fig 1 shows the time-integrated response of precipitation and atmospheric cloud radiative effect (ACRE; a positive sign indicates atmospheric heating) across horizontal grid spacing over the last 30 days, mimicking the results of Stevens and Bony (2013a). The response reveals a clear pattern as we increase the horizontal grid spacing, intensify precipitation at the ITCZ and storm tracks, decrease precipitation at the atmospheric desert, and increase ACRE, specifically in the subtropics and mid-latitudes. Beyond spatial patterns and a qualitative sense of convergence, in the following subsection, we explore relevant responses to climate, using convergence methodology as described in Peinado-Bravo et al. (2024).

B.2.1 *Canonical global mean characteristics*

Changes in the hydrological cycle are among the most known responses to climate change under increasing sea surface temperature, specifically the intensification —$\Omega_f = (f_{+4K} - f_{ctl}) \,/\, (4f_{ctl})$— of column-integrated water vapor in the atmosphere and precipitation. The Clausius-Clapeyron (CC) relationship gives us a physical constraints in the saturation vapor pressure change with temperature, often stated scaling of 6.5-7 % K^{-1} (Trenberth et al., 2003) increase of specific humidity. As the sea surface temperature increases, we expect an increase in column-integrated water vapor or precipitable water (Ω_{prw}), following CC constraint. In contrast, the increase in precipitation or intensification of the hydrological cycle (Ω_{pr}) is not constrained by the CC but rather by changes in energy availability, the net radiative cooling of the atmosphere (i.e., Allen and Ingram, 2002; Held and Soden, 2006; Jeevanjee and Romps, 2018). In our experiments, Ω_{pr} increases at a slower rate than Ω_{prw} for all horizontal resolutions (Fig B.2). At horizontal grid spacing finer than 20 km, Ω_{pr} seems insensitive to refinement with a robust convergence at 5 km horizontal grid spacing towards 3.6 $\%K^{-1}$, which is higher than Held and Soden (2006) estimate of 2-3 $\%K^{-1}$ and smaller to Guendelman et al.

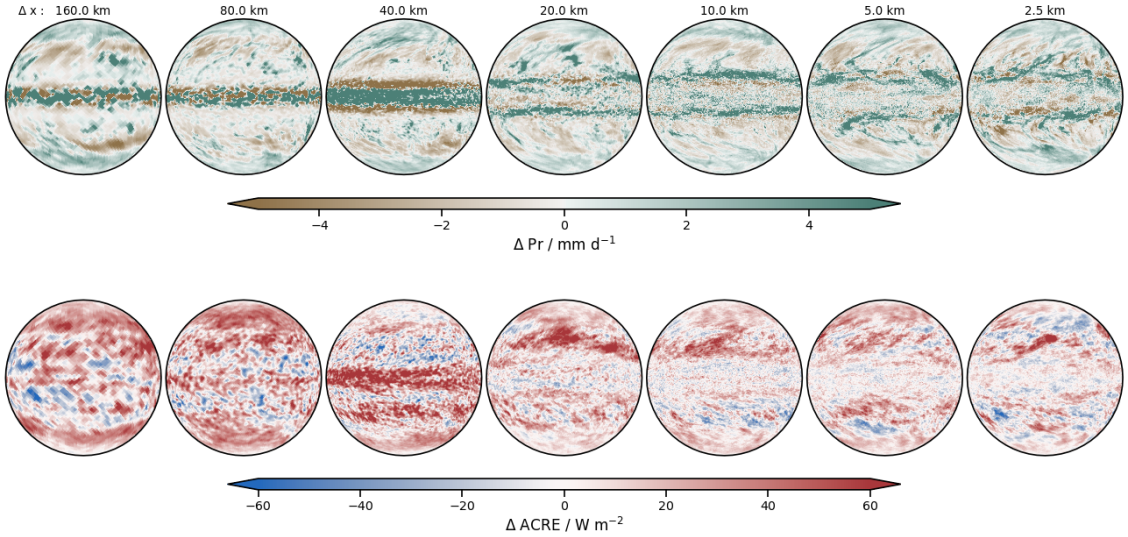


Figure B.1: Change (4k-control) in precipitation (top) and cloud radiative effect (bottom) following a uniform 4 K sea surface temperature increase for horizontal resolutions from 160 to 2.5 km averaged over the last 30 days of the aqua planet simulation. Precipitation changes show a clear signal of precipitation increase in the ITCZ with the exception of the 40 km experiment, which shows a shift an equator-ward migration, and decrease of precipitation in the atmospheric dessert. We observe a pole-ward migration of the ITCZ at finer horizontal grid spacing than 40 km. ACRE response shows a clear signal of shallow clouds increase in mid-latitudes.

(2024) estimate of 3.73 %K^{-1} over ocean. In contrast, Ω_{prw} displays a significant reduction from 20 to 10 km horizontal grid spacing, stabilizing at finer horizontal grid spacing but yet to converge at kilometer-scale horizontal grid spacing with order of convergence of 1.3 and an estimated convergence value of 10.25 %K^{-1}, larger than the CC scaling often stated.

A canonical global climate response is the change of the global mean energy balance at the top of the atmosphere (TOA)—ECS $= -\text{F}/\lambda$ — decomposed into the radiative forcing (F) and the climate feedback parameter (λ). Assuming an F of 3.7 W m^{-2}, the ECS does not display apparent convergence at 2.5 km horizontal grid spacing; instead, it has a low-reliability convergence pattern towards 1.82 K with a deviation from the estimate at 5 km horizontal grid spacing of 0.33 K (not shown). In contrast, assuming a clear sky condition, the ECS displays a high-reliable converge pattern towards 1.88 K at 5 km horizontal grid spacing, pointing out clouds as the leading source of uncertainty (not shown). In Fig B.2, we decompose the ECS into climate feedback parameters, longwave (λ_{LW}) and shortwave (λ_{LW}). λ_{LW} shows slight changes, becoming more positive as we increase the horizontal grid spacing, with a reliability convergence pattern at 5 km horizontal grid spacing to a value of -2.1 W m^{-2} K^{-1} for cloudy sky and clear sky of -2.0 W m^{-2} K^{-1} in agreement with spectral long-wave feedback parameter estimated from satellite observations $\lambda_{\text{LW}} \approx$ -2.0 W m^{-2} K^{-1} (Roemer et al., 2023). In contrast, our estimated convergence value is higher in magnitude than Merlis et al. (2024b) estimate of -1.60 W m^{-2} K^{-1} (cloudy sky), and Noda et al. (2019) estimate of -1.9 W m^{-2} K^{-1} (clear sky) and 0.22 W m^{-2} K^{-1} (cloudy sky). In contrast, λ_{SW} shows the highest sensitivity to horizontal grid spacing, going from negative to neutral, going in the opposite direction as observed by Retsch et al. (2019). λ_{SW} displays an oscillatory low-reliability convergence pattern with a linear order of convergence towards 0.14 W m^{-2} K^{-1}, similar to the muted

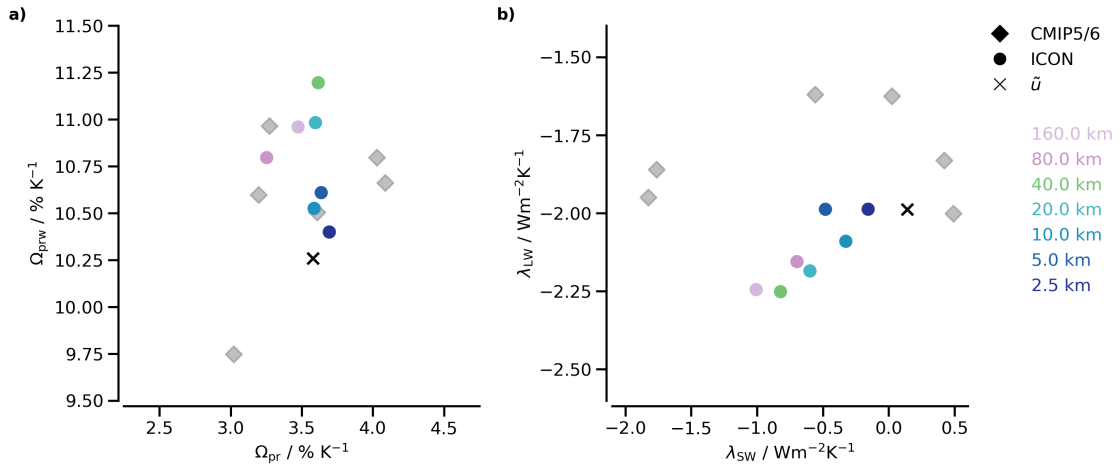


Figure B.2: The intensification of the (left) hydrological cycle (Ω_{pr}) on the x-axis versus vertical integrated water vapor (Ω_{prw}) on the y-axis and (right) climate feedback (λ) components, shortwave (x-axis) and longwave (y-axis) from 160 to 2.5 km horizontal grid spacing, in color circles, CMIP5/6 experiments, in gray diamonds, and the estimated convergence value ($\tilde{u}$), black x.

effect found by Merlis et al. (2024b) and Noda et al. (2019). The highest contributors to uncertainty in the shortwave climate feedback are the cirrus and shallow clouds in the tropical and midlatitude region – to be discussed further below.

B.2.2 *General circulation changes*

Significant changes in the large-scale circulation are expected under climate change, where different climate model projections remain uncertain (e.g., Byrne and Schneider, 2016; Byrne et al., 2018; Watt-Meyer and Frierson, 2019). The global changes in precipitation display insensitivity to the increase of horizontal grid spacing at 20 km horizontal grid spacing; however, the regional response shows disagreement as the circulation changes from one single ITCZ to double ITCZ (Figure B.3). The ITCZ location (ϕ_{pr}^{ITCZ}), defined as the precipitation maxima, displays three different responses: no movement, as a single ITCZ is well established, an equatorward movement, in the transition phase from double to single ITCZ (at 40 km horizontal grid spacing), and a slight poleward movement, as a double ITCZ is well established. From the point of view of the atmospheric energy framework, which associates the sign of the equatorial atmospheric energy input (AEI_0) with a single or double ITCZ (e.g., Adam et al., 2016; Bischoff and Schneider, 2016), we observe a positive increase of AEI_0 (1.5 Wm-2) at 40 km horizontal grid spacing associated with an equatorward ϕ_{pr}^{ITCZ} migration. As convection and deep clouds migrates with the ITCZ, we observe a considerable increase in equatorial ACRE (Fig B.1). In contrast, at finer horizontal grid spacing, the AEI_0 is more negative with a slight poleward ϕ_{pr}^{ITCZ} migration (below 1 deg). As the horizontal grid spacing reduces, the ITCZ location response displays highly reliable convergence already at 5 km horizontal grid spacing with linear order of convergence.

Along the migration of the ϕ_{pr}^{ITCZ}, we can observe the shift in precipitation, with an increase in precipitation in the deep tropics and a reduction in the atmospheric desert (ϕ^{AD}), defined as the crossing zero of the water budget (precipitation minus evaporation). In the deep tropics, the start of the atmospheric desert (ϕ_s^{AD}) and the latitude of maximum upwelling (ϕ_ψ^{ITCZ}), defined as the maxima of stream function at 500

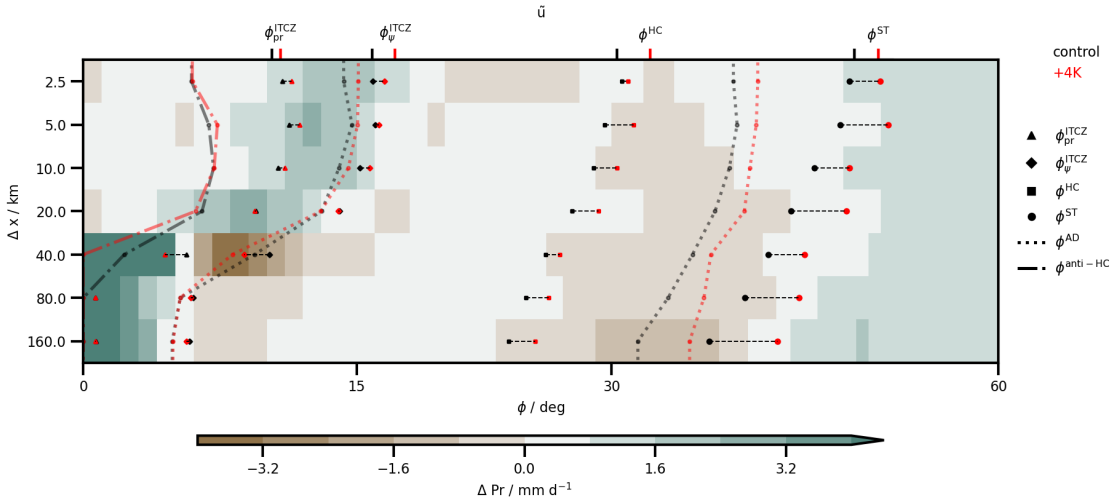


Figure B.3: The response pattern of the large-scale circulation structure, represented by the intertropical convergence zone (ITCZ), in terms of precipitation, from 160 to 2.5 km horizontal grid spacing. We define the ITCZ structure in terms of its location (maxima of precipitation: ϕ_{pr}^{ITCZ}; and maxima of the stream function at 500 hPa: ϕ_{ψ}^{ITCZ}), Hadley cell edge (ϕ^{HC}), storm track location (second maxima of precipitation; ϕ^{ST}), and atmospheric desert (ϕ^{AD}).

hPa, follow the shift of deep convection and the location of the ITCZ. Byrne et al. (2018) defined the width of the ITCZ as the ascending portion of the Hadley cell circulation. In our experiments at horizontal resolutions finer than 20 km, we observe a descending branch known as the anti-Hadley circulation (Adam et al., 2016); thus, we define the width of the ITCZ as the distance between crossing zero of the stream function at 500 hPa near the equator and ϕ_{ψ}^{ITCZ}. The ITCZ width contracts slightly (<0.2 deg) when there is a single ITCZ and expands less than 1 deg when there is a double ITCZ (not shown); differing from the systematic contraction found by Watt-Meyer and Frierson (2019).

In the sub-tropics region, we observe the poleward expansion of the Hadley cell edge (ϕ^{HC}) as observed in observation and other climate models. Watt-Meyer and Frierson (2019) observed that changes in the deep tropics, the width of the ITCZ, might lead to changes in the subtropics and extratropics, ϕ^{HC} and eddie-driven jet latitutde. At 40 km horizontal grid spacing, the expansion of ϕ^{HC} is smaller as compared to other horizontal grid spacing since the location of the tropical circulation migrated equatorward; thus, misleading at first sight the magnitude of the ϕ^{HC} expansion. By considering the changes in the ϕ_{ψ}^{ITCZ}, the width of the subtropics increases by more than 1.0 deg for all horizontal gris spacing experiments; however, it shows unclear convergence patterns as the sampling time is still too short for finer horizontal grid spacing experiments than 5 km with an estimate of 1.6 deg. Along the expansion of the Hadley cell edge, we observe the poleward movement of the atmospheric desert end, eddie-driven jet (not shown), and the storm track location (ϕ^{ST}), defined by the second maxima of precipitation. ϕ^{ST} location shifts polewards with a low-reliability convergence pattern to an estimate of 2 deg, but with uncertainty bars in the order of 1 deg due to a short sampling time.

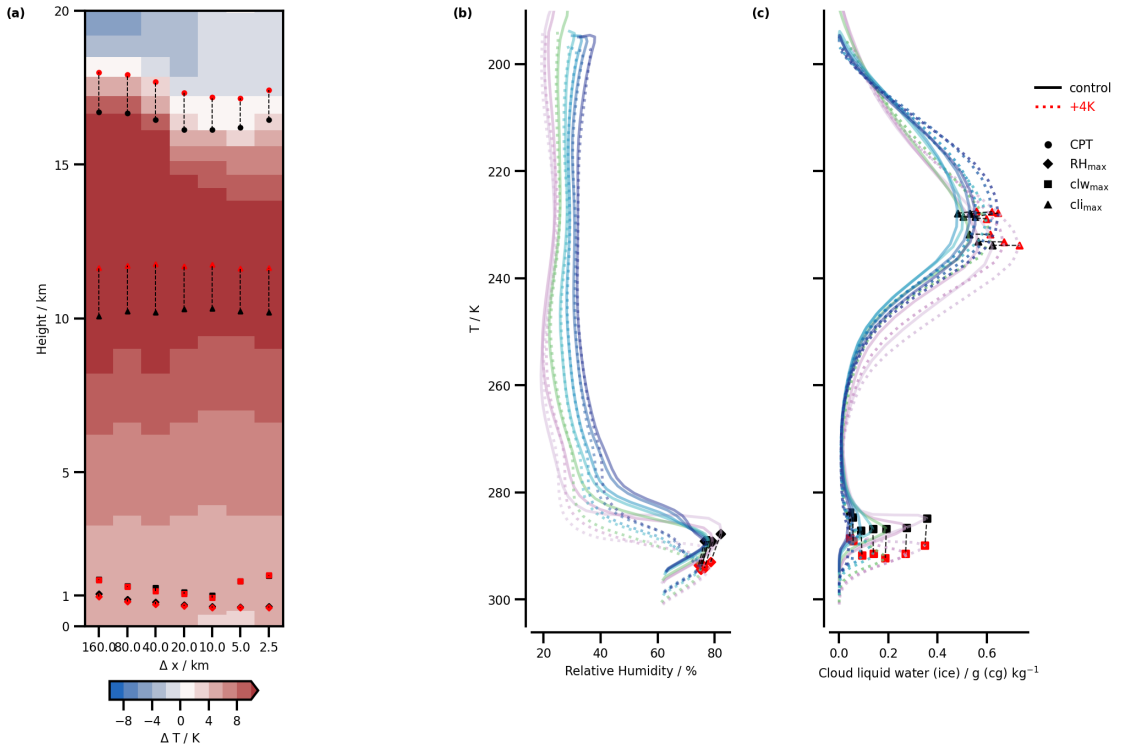


Figure B.4: Thermodynamic vertical response in the tropics (0-30 deg) of (a) temperature and (b) relative humidity in model level, and (c) cloud ice and liquid water in temperature space from 160 to 2.5 km horizontal grid spacing. We characterize the thermodynamic vertical structure regarding the cold point tropopause (CPT) and the maxima of relative humidity (RH_{max}), cloud liquid water (clw_{max}), and cloud ice cli_{max}.

B.2.3 *Thermodynamic Vertical Structure in the Tropics*

In our experiments, we observe an anti-Hadley at higher horizontal grid spacing than 20 km. To avoid the anti-Hadley cell contaminating tropical circulation response, we redefine the tropics as the region of the Hadley cell circulation, the crossing zero of the stream function near the equator to the Hadley cell edge.

Figure B.4a displays the response of the temperature vertical structure in the tropics along the location of the maxima of relative humidity (RH^{max}) and cloud liquid water in the boundary layer (clw^{max}), the cold point tropopause (CPT), and the maxima of cloud ice (cli^{max}) in the free troposphere. In agreement with other climate models, the CPT and cli^{max} shift upward with the circulation, while the boundary layer does not. The CPT has an increase of 2 K, with a low-reliability convergence pattern due to slight differences within experiments. However, we observe a substantial temperature increase in the troposphere's middle and upper levels, which agrees with different models. At 200 hPa, the temperature increases by more than a factor of 2 with respect to the forcing for all experiments. As we refine the horizontal grid spacing, the increase reduces towards an estimated convergence value of 8.4 K achieved already at 10 km horizontal grid spacing (Fig S4). In the middle of the troposphere (500 hPa), the increase in temperature is smaller than the upper levels; however, it is still more significant than the forcing of 4K in all experiments. At 500 hPa, the increase in temperature reaches convergence at 10 km horizontal grid spacing with an estimated increase of 6.5 K (Fig S4). The temperature

increase in the middle and upper levels of the troposphere is in agreement with other climate models, and compared to Merlis et al. (2024a), it is higher.

One of the biggest assumption in ECS is a constant relative humidity response (Manabe and Wetherald, 1967), approximately invariant in temperature coordinates in the free troposphere above 1 km. (Jeevanjee, 2018). Figure B.4 displays changes in relative humidity in temperature coordinates, and we can observe an evident change in the response within the boundary layer below the maxima of relative humidity and relatively no changes above it. The maxima of relative humidity in the boundary layer decreases in all experiments with changes larger than 1.5 %, displaying a reliable convergence pattern towards a 2.3 % reduction already at 20-10 km horizontal grid spacing (Fig S4). Even though the boundary layer dries, the middle and upper troposphere becomes moister at altitudes where the water vapor feedback is relevant. At pressure level of 500 hPa pressure level, near the emission temperature at clear sky condition, the relative humidity increases, with oscillatory unclear convergence pattern and large uncertanity bars, towards an estimate 2.4 % already achieved at 20 km horizontal grid spacing (Fig S4).

Figure B.4 displays the changes in cloud liquid water and cloud ice in temperature coordinates. In accordance with Hartmann and Larson (2002) and Zelinka and Hartmann (2010) hypothesis of the fixed anvil temperature (FAT) and the proportionally higher cloud base warming (PHAT), respectively, we observe small to negligible changes in the temperature profile of anvil clouds under forcing. The changes in temperature at the maxima of cloud ice are below 1K for all experiments, but small differences reduce the capability of our method to determine convergence. Closed linked to PHAT and increased stability at high altitudes with warming, Bony et al. (2016) proposed the iris effect mechanism with an expected contraction of anvil clouds. Following the convention of the International Satellite Cloud Climatology Project Simulator (Klein and Jakob, 1999; Webb et al., 2001), we define the high cloud fraction as clouds where tops are at 50-440 hPa and shallow clouds at 680-1000 hPa, using daily means of 3d cloud cover profile. The high clouds increase slightly in the tropics with a higher increase in the deep tropics, while shallow clouds in the subtropics do not show significant signs of change (Fig S1-2). However, we observe a more significant change in shallow and high cloud amounts in the extratropical region (Fig S1-2).

We explored the changes in circulation and clouds by binning the pressure vertical velocity at 500 hPa and found the probability maxima at 25 hPa d-1, in accordance to Noda et al. (2019). Low speed, near zero hPa d-1, shows an increase in updraft and downdraft events until 50 hPa d-1 and -25 hPa d-1, where events decrease significantly (Figure S6). With the increase of low-speed events, we observe a decrease in cloud cover of high clouds and a muted increase of shallow clouds at finer horizontal grid spacing except for the 5 km experiment. Higher-speed events show an increase in high cloud cover, especially in the subsidence region, and a substantial decrease in shallow clouds. The shortwave cloud radiative effect does not show significant change in the subsidence region but a moderate increase in the ascending region with a slight increase in high cloud cover.

We further investigate the impact of vertical motion and circulation changes as the grid spacing increases, comparing the change in the amount of event and the quantity at each bin of vertical velocity at 500 hPa, as performed by Terai et al. (2018) and Herrington and Reed (2020). The shortwave cloud radiative effect changes within different horizontal grid spacing are driven primarily by changes in the quantity at each bin (thermodynamic changes), reducing as we reduce the horizontal grid spacing, Figure

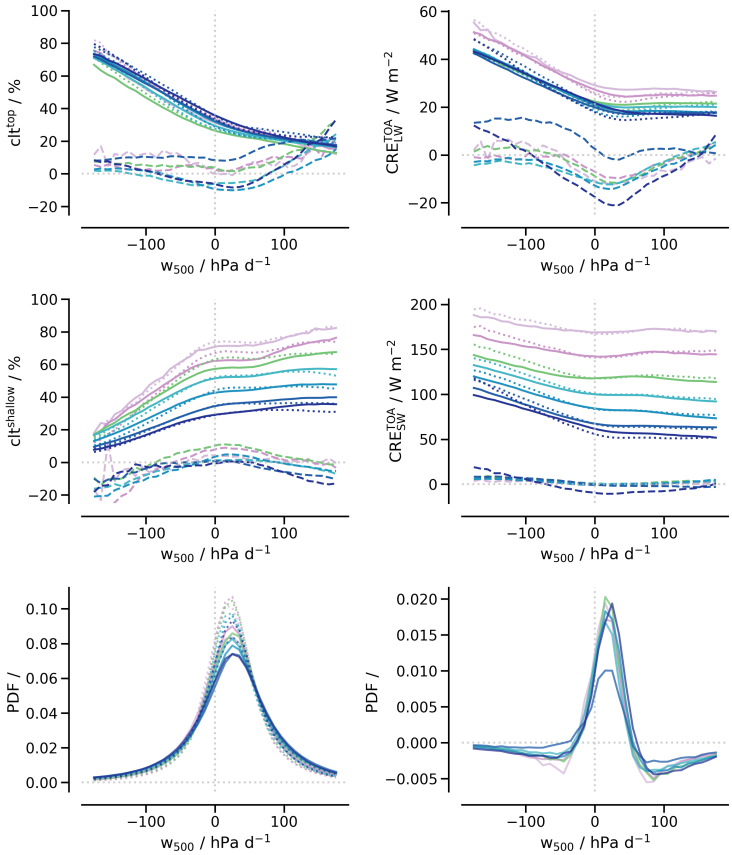


Figure B.5: Cloud fraction and cloud radiative effect as a function of vertical velocity at 500 hPa (w_{500}) in the tropics defined as the region as the region of the Hadley cell circulation, the crossing zero of the stream function near the equator to the Hadley cell edge to avoid the anti-Hadley cell. Positive w_{500} indicates downdraft and negative updrafts. (a) High clouds defined as clouds with cloud top height at 50-440hPa, (c) shallow clouds at 680-1,000 hPa, following the International Satellite Cloud Climatology project simulator. (b) Longwave cloud radiative effect (warming) and (d) shortwave cloud radiative effect (cooling). (e) The probability distribution of w_{500} and (f) the difference within forced and control simulation of the probability distribution of w_{500}. Solid lines indicates control simulation, dotted lines simulation with an increase of SST +4K (forced), and dash lines the relative changes with respect of control. Color labels indicates the horizontal grid spacing of each experiment.

S9. As we increase the horizontal grid spacing, the changes in circulation within different horizontal grid spacing converge in response to forcing. In contrast, the longwave cloud radiative effect changes within different horizontal grid spacing display similar responses for change in the amount of event (dynamic changes) and the quantity per bin (thermodynamic changes), Figure S8. Thus, it presents relatively small difference and insensitivity to horizontal grid spacing increase, even in the response under forcing. Meanwhile, by decomposing the dynamic and thermodynamic changes under the response to forcing, following Bony and Dufresne (2005) methodology, the shortwave and longwave cloud radiative effect has an increasing effect by the dynamical response, while the thermodynamic response reduces its effect (Figure S8-9). In both cases, the dynamical response is stronger than the thermodynamic response, disagreeing with the finding of Noda et al. (2019) in the shortwave cloud radiative effect.

B.3 SUMMARY AND CONCLUSIONS

Using a global storm resolving model (GSRM), ICON, we have investigated atmosphere's responses without convection parameterization under forcing, mimicking climate change, across a hierarchy of resolutions from 160 to 1.25 km horizontal grid spacing. We employed an aquaplanet experimental setup to reduce the computational constraints of high-resolution experiments to achieve a statistically steady state and to explore if, by increasing the horizontal grid spacing, the atmosphere displays physical convergence or how sensible the response to resolution is. Our main findings are:

- Precipitation intensification (Ω_{pr}) converges at 20-10 km horizontal resolution with an estimated value of 3.6 % K^{-1}. In contrast, the increase of column-integrated water vapor (Ω_{prw}) does not converge at a kilometer-scale with a shift at 10 km horizontal grid spacing with a Ω_{prw} than CC-scaling.
- Longwave climate feedback (λ_{LW}) does not show significant changes with increasing horizontal grid spacing and converges at kilometer-scale horizontal grid spacing (5 km) compared to the shortwave component (λ_{SW}), which requires finer horizontal grid spacing to converge towards an approximately neutral feedback.
- Circulation response is affected if the control state is a single, double, or unstable ITCZ, with non, poleward, or equatorward migration of its location ($\phi_{\psi}^{\mathrm{ITCZ}}$ and $\phi_{\mathrm{pr}}^{\mathrm{ITCZ}}$) and contraction, expansion, or expansion of its width.
- For any control state, the Hadley cell edge (ϕ^{HC}) has a poleward migration along the storm track latitude (ϕ^{ST}) with apparent/oscillatory convergence at 10-5 km horizontal grid spacing. However, due to the short sampling time, they present large uncertainty bars obscuring their convergence.
- For all experiments, relative humidity decreases near the boundary layer, with its maxima converging already at 20-10 km horizontal grid spacing.
- Above the boundary layer, the changes in relative humidity are close to zero; however, at heights near clear-sky emission temperature, it has a positive sign with apparent convergence at 10 km horizontal grid spacing.
- High clouds in the tropics shift upwards, nearly following an isothermal (FAT and PHAT hypothesis). However, we observe a slight increase of high clouds in deep tropics, in opposite expectation of the iris effect.
- We do not observe a significant increase of shallow clouds in the descending branch of the Hadley cell; instead, we observe an almost neutral response. However, the systematic reduction of shallow clouds in the tropics might obscure this response.

With the increase of horizontal grid spacing and the avoidance of convection parameterization, we expected a robust response in diverse climate responses/feedbacks in a warming climate. In the global mean, we observed a robust Ω_{pr}, which agrees with the findings of the X-Shield AMIP simulations of 3.73 % K^{-1} over the ocean and within the CMIP5/6 estimate (Guendelman et al., 2024). Meanwhile, the λ_{LW} estimate agrees with observational findings (Roemer et al., 2023) but is higher than other AMIP experiments using GSRMs (Noda et al., 2019; Merlis et al., 2024b), while the shortwave component agrees well with the neutral feedback observed by Merlis et al. (2024b). The circulation

response shows an apparent and qualitative convergence at horizontal grid spacing finer than 10 km, but our methodology and sampling time do not observe a robust convergence pattern. However, the effects of the reference climate state and changes in tropical circulation affect the response of subtropics and midlatitude metrics, which aligns with Watt-Meyer and Frierson (2019) findings.

An intriguing result is the robust reduction and increase of relative humidity in the boundary layer and at clear-sky emission height. The changes in relative humidity at clear-sky emission height play a significant role in the longwave climate feedback (Retsch et al., 2019), which is robust across horizontal grid spacing. Thus, it poses the question of how large the spread in the response is within other GSRMs and could lead to explain the inter-model differences in the λ_{LW}. We remain positive in this aspect, as Lang et al. (2021) observed small relative humidity intermodel differences within DYAMOND (Stevens et al., 2019), and Bourdin et al. (2021) suggested that water-vapor-lapse rate feedback depends on its initial relative humidity.

Another key result is the non-observed iris effect in the deep tropics, which has been observed in GCMs and GSRMS within radiative convective equilibrium experiments (Silvers et al., 2023). Jeevanjee (2022) explored the possibility that the mass convective flux closure base for the iris effect hypothesis (Bony et al., 2016) is only partially accurate since relative humidity and precipitation efficiency do not remain constant. We observe changes in relative humidity, which might be significant for the iris effect hypothesis. On the other hand, we observe a neutral response in the shallow clouds in the tropics, but there are significant changes in the midlatitudes, which drives differences in λ.

We have shown that the atmosphere component under a climate change scenario displays physical convergence and agreement with other models. However, many questions still need to be answered regarding the observed difference, such as why the iris effect is not observed in our experiments or other GSRMs? How do the changes in relative humidity play a role, and how significant are those differences? The upcoming third phase of DYAMOND aims to provide a more extensive and complete data set and analysis to answer these questions using a hierarchy of experiments, such as aquaplanet experiments (ImDYAMOND3) and long time integrated AMIP simulations.

B.4 SIMULATION AND METHODS

We perform a series of global aquaplanet experiments with horizontal grid spacing ranging from 160 km to 1.25km as described in (Peinado-Bravo et al., 2024) using the atmosphere component of ICON (version 2.6.5[1]). We retain only parameterizations for the radiant energy transfer, which is based on RRTMGP (Rapid Radiative Transfer model for General circulation model applications, Parallel, Pincus et al., 2019), a single moment bulk microphysical parameterization consisting of five condensate habits (Baldauf et al., 2011), and 3D turbulent mixing as described in Smagorinsky (1963) with modifications by Lilly (1962) as implemented in the ICON model by Lee et al. (2022b) following an earlier implementation by Dipankar et al. (2015). Further details of the present configuration are provided by Hohenegger et al. (2022).

Our control experiments follow the protocol of the AquaPlanet Experiment (Blackburn et al., 2013), consist of an Earth-sized planet with an Earth-like atmosphere whose lower boundary condition is consistent with a water (no sea-ice) covered surface with a prescribed and zonally and temporally constant sea surface temperature (T_s) (Neale and

[1] icon/icon-2.6.5-rc - hash f6b6ed9f1

Hoskins, 2000). In addition, it maintains a perpetual equinox with symmetric-constant irradiation about the equator by setting the eccentricity and obliquity of earth to zero, and using a solar constant of 1361 W/m^2. Ozone follows a constant zonally symmetric profile with respect to the equator, the greenhouse gases are well mixed, and the interaction between aerosol and radiation is neglected. Meanwhile, our forced experiments have a constant increase of +4K in the SST only.

For our analysis we extend the simulation performed by Peinado-Bravo et al. (2024) by 1 year, 180 days and 90 days for experiments with horizontal grid spacing ranging from 160-80, 40-5, and 2.5 km respectively for our control and forced scenarios. We discard the first 45 days of forced experiments as the spin-up time. We do not use the 1.25 km horizontal grid spacing experiment due to computational constraints. [Healpis output? mention?]. For comparison with CMIPs aquaplanet experiments we use the aqua-control and aqua-p4K experiments of CMIP5 (CNRM-CM5, HadGEM2A, IPSL-CM5A-LR, MPI-EMS-LR, MPI-EMS-MR) and CMIP6 (CESM2, CNRM-CM6-1, GFDL, IPSL-LR).

OPEN RESEARCH SECTION

The aquaplanet experiments were done with the ICON release candidate 2.6.5 as commit f6b6ed9f1. The source code is available at https://doi.org/10.17617/3.1XTSR6, and the ICON model is available to individuals under licenses https://mpimet.mpg.de/en/science/modeling-with-icon/code-availability.

ACKNOWLEDGEMENTS

The authors would like to thank C. Hohenegger, H. Segura, and J. Bao for valuable discussion and feedback We are grateful for advice and technical support from colleagues at DKRZ and MPI-M, specifically from R. Redler and JF. Engels. The research was supported by public funding to the Max Planck Society, the European Union's Horizon 2020 research and innovation program project NextGEMS funded C. Hohenegger and B. Stevens under the grant agreement number 101003470, and the WarmWorld funded by the German Federal Ministry of Education and Research under the funding code 01LK2202B.

BIBLIOGRAPHY

Adam, O., K. M. Grise, P. Staten, I. R. Simpson, S. M. Davis, N. A. Davis, D. W. Waugh, T. Birner, and A. Ming (2018). "The TropD software package (v1): standardized methods for calculating tropical-width diagnostics". *Geoscientific Model Development* 11.10, pp. 4339–4357. DOI: 10.5194/gmd-11-4339-2018.

Adam, Ori (2021). "Dynamic and Energetic Constraints on the Modality and Position of the Intertropical Convergence Zone in an Aquaplanet". *Journal of Climate* 34.2, pp. 527–543. DOI: 10.1175/JCLI-D-20-0128.1.

Adam, Ori, Tobias Bischoff, and Tapio Schneider (2016). "Seasonal and Interannual Variations of the Energy Flux Equator and ITCZ. Part II: Zonally Varying Shifts of the ITCZ". *Journal of Climate* 29.20, pp. 7281–7293. DOI: 10.1175/JCLI-D-15-0710.1.

Allen, Myles R. and William J. Ingram (Sept. 2002). "Constraints on future changes in climate and the hydrologic cycle". *Nature* 419.6903, pp. 224–232. ISSN: 1476-4687. DOI: 10.1038/nature01092.

Baker, Alexander J., Benoît Vannière, and Pier Luigi Vidale (2024). "On the Realism of Tropical Cyclone Intensification in Global Storm-Resolving Climate Models". *Geophysical Research Letters* 51.17, e2024GL109841. DOI: https://doi.org/10.1029/2024GL109841.

Baldauf, Michael, Axel Seifert, Jochen Förstner, Detlev Majewski, Matthias Raschendorfer, and Thorsten Reinhardt (2011). "Operational Convective-Scale Numerical Weather Prediction with the COSMO Model: Description and Sensitivities". *Monthly Weather Review* 139.12, pp. 3887–3905. DOI: 10.1175/MWR-D-10-05013.1.

Bechtold, Peter, Irina Sandu, D Klocke, N Semane, Maike Ahlgrimm, Anton Beljaars, Richard Forbes, and Mark Rodwell (2014). *The role of shallow convection in ECMWF's Integrated Forecasting System*.

Bischoff, Tobias and Tapio Schneider (2016). "The Equatorial Energy Balance, ITCZ Position, and Double-ITCZ Bifurcations". *Journal of Climate* 29.8, pp. 2997–3013. DOI: 10.1175/JCLI-D-15-0328.1.

Blackburn, Michael, David L Williamson, Kensuke Nakajima, Wataru Ohfuchi, Yoshiyuki O Takahashi, Yoshi-Yuki Hayashi, et al. (2013). "The Aqua-Planet Experiment (APE): CONTROL SST Simulation". *Journal of the Meteorological Society of Japan. Ser. II* 91A, pp. 17–56. DOI: 10.2151/jmsj.2013-A02.

Boer, George J. and B. Denis (1997). "Numerical convergence of the dynamics of a GCM". *Climate Dynamics* 13, pp. 359–374. URL: https://api.semanticscholar.org/CorpusID:128428396.

Bogenschutz, Peter A. and Steven K. Krueger (2013). "A simplified PDF parameterization of subgrid-scale clouds and turbulence for cloud-resolving models". *Journal of Advances in Modeling Earth Systems* 5.2, pp. 195–211. DOI: https://doi.org/10.1002/jame.20018.

Bony, Sandrine and Jean-Louis Dufresne (2005). "Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models". *Geophysical Research Letters* 32.20. DOI: https://doi.org/10.1029/2005GL023851.

Bony, Sandrine, Bjorn Stevens, David Coppin, Tobias Becker, Kevin A. Reed, Aiko Voigt, and Brian Medeiros (2016). "Thermodynamic control of anvil cloud amount".

Proceedings of the National Academy of Sciences 113.32, pp. 8927–8932. DOI: 10.1073/pnas.1601472113.

Bony, Sandrine, Bjorn Stevens, Dargan M. W. Frierson, Christian Jakob, Masa Kageyama, Robert Pincus, et al. (Apr. 2015). "Clouds, circulation and climate sensitivity". *Nature Geoscience* 8.4, pp. 261–268. ISSN: 1752-0908. DOI: 10.1038/ngeo2398.

Bourdin, Stella, Lukas Kluft, and Bjorn Stevens (2021). "Dependence of Climate Sensitivity on the Given Distribution of Relative Humidity". *Geophysical Research Letters* 48.8, e2021GL092462. DOI: https://doi.org/10.1029/2021GL092462.

Byrne, Michael P., Angeline G. Pendergrass, Anita D. Rapp, and Kyle R. Wodzicki (Dec. 2018). "Response of the Intertropical Convergence Zone to Climate Change: Location, Width, and Strength". *Current Climate Change Reports* 4.4, pp. 355–370. ISSN: 2198-6061. DOI: 10.1007/s40641-018-0110-5.

Byrne, Michael P. and Tapio Schneider (2016). "Narrowing of the ITCZ in a warming climate: Physical mechanisms". *Geophysical Research Letters* 43.21, pp. 11, 350–11, 357. DOI: https://doi.org/10.1002/2016GL070396.

Ceppi, Paulo and Dennis L. Hartmann (2013). "On the Speed of the Eddy-Driven Jet and the Width of the Hadley Cell in the Southern Hemisphere". *Journal of Climate* 26.10, pp. 3450–3465. DOI: 10.1175/JCLI-D-12-00414.1.

Chahine, Moustafa T. (Oct. 1992). "The hydrological cycle and its influence on climate". *Nature* 359.6394, pp. 373–380. ISSN: 1476-4687. DOI: 10.1038/359373a0.

Charney, J. G. (1979). "Carbon Dioxide and Climate: A Scientific Assessment", p. 33.

Chen, Liang and Oliver W. Frauenfeld (2014). "Surface Air Temperature Changes over the Twentieth and Twenty-First Centuries in China Simulated by 20 CMIP5 Models". *Journal of Climate* 27.11, pp. 3920–3937. DOI: 10.1175/JCLI-D-13-00465.1.

Chinita, M. J., M. Witte, M. J. Kurowski, J. Teixeira, K. Suselj, G. Matheou, and P. Bogenschutz (2023). "Improving the representation of shallow cumulus convection with the simplified-higher-order-closure–mass-flux (SHOC+MF v1.0) approach". *Geoscientific Model Development* 16.7, pp. 1909–1924. DOI: 10.5194/gmd-16-1909-2023.

Chou, Chia and J. David Neelin (2004). "Mechanisms of Global Warming Impacts on Regional Tropical Precipitation". *Journal of Climate* 17.13, pp. 2688–2701. DOI: 10.1175/1520-0442(2004)017<2688:MOGWIO>2.0.CO;2.

Chou, Chia, J. David Neelin, Chao-An Chen, and Jien-Yi Tu (2009). "Evaluating the "Rich-Get-Richer" Mechanism in Tropical Precipitation Change under Global Warming". *Journal of Climate* 22.8, pp. 1982–2005. DOI: 10.1175/2008JCLI2471.1.

Dipankar, Anurag, Bjorn Stevens, Rieke Heinze, Christopher Moseley, Günther Zängl, Marco Giorgetta, and Slavko Brdar (2015). "Large eddy simulation using the general circulation model ICON". *Journal of Advances in Modeling Earth Systems* 7.3, pp. 963–986. DOI: 10.1002/2015MS000431.

Douville, Hervé, Saïd Qasmi, Aurélien Ribes, and Olivier Bock (Oct. 2022). "Global warming at near-constant tropospheric relative humidity is supported by observations". *Communications Earth & Environment* 3.1, p. 237. ISSN: 2662-4435. DOI: 10.1038/s43247-022-00561-z.

Eliassen, A. (1948). "The quasi-static equations of motion." *Geofys. Publ.* 17.

Feng, Zhe, L. Ruby Leung, Joseph Hardin, Christopher R. Terai, Fengfei Song, and Peter Caldwell (2023). "Mesoscale Convective Systems in DYAMOND Global Convection-Permitting Simulations". *Geophysical Research Letters* 50.4, e2022GL102603. DOI: https://doi.org/10.1029/2022GL102603.

Freischem, Lilli J., Philipp Weiss, Hannah M. Christensen, and Philip Stier (2024). "Multifractal Analysis for Evaluating the Representation of Clouds in Global Kilometer-

Scale Models". *Geophysical Research Letters* 51.20, e2024GL110124. DOI: https://doi.org/10.1029/2024GL110124.

Gassmann, Almut and Hans-Joachim Herzog (2008). "Towards a consistent numerical compressible non-hydrostatic model using generalized Hamiltonian tools". *Quarterly Journal of the Royal Meteorological Society* 134.635, pp. 1597–1613. DOI: 10.1002/qj.297.

Giorgetta, M. A., R. Brokopf, T. Crueger, M. Esch, S. Fiedler, J. Helmert, et al. (2018). "ICON-A, the Atmosphere Component of the ICON Earth System Model: I. Model Description". *Journal of Advances in Modeling Earth Systems* 10.7, pp. 1613–1637. DOI: 10.1029/2017MS001242.

Giorgetta, M. A., W. Sawyer, X. Lapillonne, P. Adamidis, D. Alexeev, V. Clément, et al. (2022). "The ICON-A model for direct QBO simulations on GPUs (version icon-cscs:baf28a514)". *Geoscientific Model Development* 15.18, pp. 6985–7016. DOI: 10.5194/gmd-15-6985-2022.

Guendelman, Ilai, Timothy M. Merlis, Kai-Yuan Cheng, Lucas M. Harris, Christopher S. Bretherton, Maximilien Bolot, Linjiong Zhou, Alex Kaltenbaugh, Spencer K. Clark, and Stephan Fueglistaler (2024). "The Precipitation Response to Warming and CO2 Increase: A Comparison of a Global Storm Resolving Model and CMIP6 Models". *Geophysical Research Letters* 51.7, e2023GL107008. DOI: https://doi.org/10.1029/2023GL107008.

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

Held, Isaac M. (2005). "The Gap between Simulation and Understanding in Climate Modeling". *Bulletin of the American Meteorological Society* 86.11, pp. 1609–1614. DOI: 10.1175/BAMS-86-11-1609.

Held, Isaac M. and Brian J. Soden (2006). "Robust Responses of the Hydrological Cycle to Global Warming". *Journal of Climate* 19.21, pp. 5686–5699. DOI: 10.1175/JCLI3990.1.

Herrington, Adam R. and Kevin A. Reed (2020). "On resolution sensitivity in the Community Atmosphere Model". *Quarterly Journal of the Royal Meteorological Society* 146.733, pp. 3789–3807. DOI: https://doi.org/10.1002/qj.3873.

Hohenegger, C., P. Korn, L. Linardakis, R. Redler, R. Schnur, P. Adamidis, et al. (2022). "ICON-Sapphire: simulating the components of the Earth System and their interactions at kilometer and subkilometer scales". *Geoscientific Model Development Discussions* 2022, pp. 1–42. DOI: 10.5194/gmd-2022-171.

Hohenegger, Cathy, Luis Kornblueh, Daniel Klocke, Tobias Becker, Guido Cioni, Jan Frederik Engels, Uwe Schulzweida, and Bjorn Stevens (2020). "Climate Statistics in Global Simulations of the Atmosphere, from 80 to 2.5 km Grid Spacing". *Journal of the Meteorological Society of Japan. Ser. II* 98.1, pp. 73–91. DOI: 10.2151/jmsj.2020-005.

Hughes, Ifan G. and Thomas P. A. Hase (May 2012). "Error Propagation: A Functional Approach". *Journal of Chemical Education* 89.6, pp. 821–822. ISSN: 0021-9584. DOI: 10.1021/ed2004627.

J. Freitas, Dr. Christopher (Sept. 1993). "Editorial". *Journal of Fluids Engineering* 115.3, pp. 339–340. ISSN: 0098-2202. DOI: 10.1115/1.2910144.

J. G. Charney, R. FjÖrtoft and J. Von Neumann (1950). "Numerical Integration of the Barotropic Vorticity Equation". *Tellus* 2.4, pp. 237–254. DOI: 10.3402/tellusa.v2i4.8607.

Jeevanjee, Nadir (Feb. 2018). "The physics of climate change: simple models in climate science".

Jeevanjee, Nadir (2022). "Three Rules for the Decrease of Tropical Convection With Global Warming". *Journal of Advances in Modeling Earth Systems* 14.11. e2022MS003285 2022MS003285, e2022MS003285. DOI: https://doi.org/10.1029/2022MS003285.

Jeevanjee, Nadir, Pedram Hassanzadeh, Spencer Hill, and Aditi Sheshadri (2017). "A perspective on climate model hierarchies". *Journal of Advances in Modeling Earth Systems* 9.4, pp. 1760–1771. DOI: https://doi.org/10.1002/2017MS001038.

Jeevanjee, Nadir and David M. Romps (2018). "Mean precipitation change from a deepening troposphere". *Proceedings of the National Academy of Sciences* 115.45, pp. 11465–11470. DOI: 10.1073/pnas.1720683115.

Judt, Falko, Daniel Klocke, Rosimar Rios-Berrios, Benoit Vanniere, Florian Ziemen, Ludovic Auger, et al. (2021). "Tropical Cyclones in Global Storm-Resolving Models". *Journal of the Meteorological Society of Japan. Ser. II* 99.3, pp. 579–602. DOI: 10.2151/jmsj.2021-029.

Kajikawa, Yoshiyuki, Yoshiaki Miyamoto, Ryuji Yoshida, Tsuyoshi Yamaura, Hisashi Yashiro, and Hirofumi Tomita (June 2016). "Resolution dependence of deep convections in a global simulation from over 10-kilometer to sub-kilometer grid spacing". *Progress in Earth and Planetary Science* 3.1, p. 16. ISSN: 2197-4284. DOI: 10.1186/s40645-016-0094-5.

Kang, Sarah M. and Lorenzo M. Polvani (2011). "The Interannual Relationship between the Latitude of the Eddy-Driven Jet and the Edge of the Hadley Cell". *Journal of Climate* 24.2, pp. 563–568. DOI: 10.1175/2010JCLI4077.1.

Klein, Stephen A. and Christian Jakob (1999). "Validation and Sensitivities of Frontal Clouds Simulated by the ECMWF Model". *Monthly Weather Review* 127.10, pp. 2514–2531. DOI: 10.1175/1520-0493(1999)127<2514:VASOFC>2.0.CO;2.

Klocke, Daniel, Robert Pincus, and Johannes Quaas (2011). "On Constraining Estimates of Climate Sensitivity with Present-Day Observations through Model Weighting". *Journal of Climate* 24.23, pp. 6092–6099. DOI: 10.1175/2011JCLI4193.1.

Kodama, Chihiro, Yohei Yamada, Akira T. Noda, Kazuyoshi Kikuchi, Yoshiyuki Kajikawa, Tomoe Nasuno, et al. (2015). "A 20-Year Climatology of a NICAM AMIP-Type Simulation". *Journal of the Meteorological Society of Japan. Ser. II* 93.4, pp. 393–424. DOI: 10.2151/jmsj.2015-024.

Koutroulis, A. G., M. G. Grillakis, I. K. Tsanis, and L. Papadimitriou (Nov. 2016). "Evaluation of precipitation and temperature simulation performance of the CMIP3 and CMIP5 historical experiments". *Climate Dynamics* 47.5, pp. 1881–1898. ISSN: 1432-0894. DOI: 10.1007/s00382-015-2938-x.

Landu, Kiranmayi, L. Ruby Leung, Samson Hagos, V. Vinoj, Sara A. Rauscher, Todd Ringler, and Mark Taylor (2014). "The Dependence of ITCZ Structure on Model Resolution and Dynamical Core in Aquaplanet Simulations". *Journal of Climate* 27.6, pp. 2375–2385. DOI: 10.1175/JCLI-D-13-00269.1.

Lang, Theresa, Ann Kristin Naumann, Bjorn Stevens, and Stefan A. Buehler (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* 13.11, e2021MS002514. DOI: https://doi.org/10.1029/2021MS002514.

Langhans, Wolfgang, Juerg Schmidli, and Christoph Schär (2012). "Bulk Convergence of Cloud-Resolving Simulations of Moist Convection over Complex Terrain". *Journal of the Atmospheric Sciences* 69.7, pp. 2207–2228. DOI: 10.1175/JAS-D-11-0252.1.

Lee, Junhong, Cathy Hohenegger, Andreas Chlond, and Reiner Schnur (June 2022a). "The Climatic Role of Interactive Leaf Phenology in the Vegetation- Atmosphere

System of Radiative-Convective Equilibrium Storm-Resolving Simulations". *Tellus B: Chemical and Physical Meteorology*. DOI: 10.16993/tellusb.26.

– (June 2022b). "The Climatic Role of Interactive Leaf Phenology in the Vegetation- Atmosphere System of Radiative-Convective Equilibrium Storm-Resolving Simulations". *Tellus B: Chemical and Physical Meteorology*. DOI: 10.16993/tellusb.26.

Lilly, Douglas K (1962). "On the numerical simulation of buoyant convection". *Tellus* 14.2, pp. 148–172.

Lindzen, Richard S., Ming-Dah Chou, and Arthur Y. Hou (2001). "Does the Earth Have an Adaptive Infrared Iris?" *Bulletin of the American Meteorological Society* 82.3, pp. 417–432. DOI: 10.1175/1520-0477(2001)082<0417:DTEHAA>2.3.CO;2.

Liu, J., M. Song, Y. Hu, and X. Ren (2012). "Changes in the strength and width of the Hadley Circulation since 1871". *Climate of the Past* 8.4, pp. 1169–1175. DOI: 10.5194/cp-8-1169-2012.

Lu, Jian, Gang Chen, L. Ruby Leung, D. Alex Burrows, Qing Yang, Koichi Sakaguchi, and Samson Hagos (2015). "Toward the Dynamical Convergence on the Jet Stream in Aquaplanet AGCMs". *Journal of Climate* 28.17, pp. 6763–6782. DOI: 10.1175/JCLI-D-14-00761.1.

Lu, Jian, Gabriel A. Vecchi, and Thomas Reichler (2007). "Expansion of the Hadley cell under global warming". *Geophysical Research Letters* 34.6. DOI: https://doi.org/10.1029/2006GL028443.

Lynch, Peter (1992). "Richardson's Barotropic Forecast: A Reappraisal". *Bulletin of the American Meteorological Society* 73.1, pp. 35–48. DOI: 10.1175/1520-0477(1992)073<0035:RBFAR>2.0.CO;2.

Ma, Hsi-Yen, Stephen A. Klein, Jiwoo Lee, Min-Seop Ahn, Cheng Tao, and Peter J. Gleckler (2022). "Superior Daily and Sub-Daily Precipitation Statistics for Intense and Long-Lived Storms in Global Storm-Resolving Models". *Geophysical Research Letters* 49.8, e2021GL096759. DOI: https://doi.org/10.1029/2021GL096759.

Manabe, Syukuro and Kirk Bryan (1969). "Climate Calculations with a Combined Ocean-Atmosphere Model". *Journal of Atmospheric Sciences* 26.4, pp. 786–789. DOI: 10.1175/1520-0469(1969)026<0786:CCWACO>2.0.CO;2.

Manabe, Syukuro and Richard T. Wetherald (1967). "Thermal Equilibrium of the Atmosphere with a Given Distribution of Relative Humidity". *Journal of Atmospheric Sciences* 24.3, pp. 241–259. DOI: 10.1175/1520-0469(1967)024<0241:TEOTAW>2.0.CO;2.

Mbengue, Cheikh and Tapio Schneider (2013). "Storm Track Shifts under Climate Change: What Can Be Learned from Large-Scale Dry Dynamics". *Journal of Climate* 26.24, pp. 9923–9930. DOI: 10.1175/JCLI-D-13-00404.1.

– (2018). "Linking Hadley Circulation and Storm Tracks in a Conceptual Model of the Atmospheric Energy Balance". *Journal of the Atmospheric Sciences* 75.3, pp. 841–856. DOI: 10.1175/JAS-D-17-0098.1.

Medeiros, Brian, Bjorn Stevens, and Sandrine Bony (Apr. 2015). "Using aquaplanets to understand the robust responses of comprehensive climate models to forcing". *Climate Dynamics* 44.7, pp. 1957–1977. ISSN: 1432-0894. DOI: 10.1007/s00382-014-2138-0.

Medeiros, Brian, David L. Williamson, and Jerry G. Olson (2016). "Reference aquaplanet climate in the Community Atmosphere Model, Version 5". *Journal of Advances in Modeling Earth Systems* 8.1, pp. 406–424. DOI: https://doi.org/10.1002/2015MS000593.

Meehl, Gerald A., Catherine A. Senior, Veronika Eyring, Gregory Flato, Jean-Francois Lamarque, Ronald J. Stouffer, Karl E. Taylor, and Manuel Schlund (2020). "Context for interpreting equilibrium climate sensitivity and transient climate response from the

CMIP6 Earth system models". *Science Advances* 6.26, eaba1981. DOI: 10.1126/sciadv.aba1981.

Merlis, Timothy M, Ilai Guendelman, Kai-Yuan Cheng, Lucas Harris, Yan-Ting Chen, Christopher S Bretherton, et al. (July 2024a). "The vertical structure of tropical temperature change in global storm-resolving model simulations of climate change".

Merlis, Timothy M., Kai-Yuan Cheng, Ilai Guendelman, Lucas Harris, Christopher S. Bretherton, Maximilien Bolot, et al. (2024b). "Climate sensitivity and relative humidity changes in global storm-resolving model simulations of climate change". *Science Advances* 10.26, eadn5217. DOI: 10.1126/sciadv.adn5217.

Miura, Hiroaki, Masaki Satoh, Tomoe Nasuno, Akira T. Noda, and Kazuyoshi Oouchi (2007a). "A Madden-Julian Oscillation Event Realistically Simulated by a Global Cloud-Resolving Model". *Science* 318.5857, pp. 1763–1765. DOI: 10.1126/science.1148443.

Miura, Hiroaki, Masaki Satoh, Hirofumi Tomita, Akira T. Noda, Tomoe Nasuno, and Shin-ichi Iga (2007b). "A short-duration global cloud-resolving simulation with a realistic land and sea distribution". *Geophysical Research Letters* 34.2. DOI: 10.1029/2006GL027448.

Miyakawa, Tomoki, Masaki Satoh, Hiroaki Miura, Hirofumi Tomita, Hisashi Yashiro, Akira T. Noda, Yohei Yamada, Chihiro Kodama, Masahide Kimoto, and Kunio Yoneyama (May 2014). "Madden–Julian Oscillation prediction skill of a new-generation global model demonstrated using a supercomputer". *Nature Communications* 5.1, p. 3769. ISSN: 2041-1723. DOI: 10.1038/ncomms4769.

Möbis, Benjamin and Bjorn Stevens (2012). "Factors controlling the position of the Intertropical Convergence Zone on an aquaplanet". *Journal of Advances in Modeling Earth Systems* 4.4. DOI: https://doi.org/10.1029/2012MS000199.

Narenpitak, Pornampai, Christopher S. Bretherton, and Marat F. Khairoutdinov (2017). "Cloud and circulation feedbacks in a near-global aquaplanet cloud-resolving model". *Journal of Advances in Modeling Earth Systems* 9.2, pp. 1069–1090. DOI: https://doi.org/10.1002/2016MS000872.

Neale, R. B. and B. J. Hoskins (2000). "A standard test for AGCMs including their physical parametrizations: I: the proposal". *Atmospheric Science Letters* 1.2, pp. 101–107. DOI: https://doi.org/10.1006/asle.2000.0022.

Noda, Akira T., Chihiro Kodama, Yohei Yamada, Masaki Satoh, Tomoo Ogura, and Tomoki Ohno (2019). "Responses of Clouds and Large-Scale Circulation to Global Warming Evaluated From Multidecadal Simulations Using a Global Nonhydrostatic Model". *Journal of Advances in Modeling Earth Systems* 11.9, pp. 2980–2995. DOI: https://doi.org/10.1029/2019MS001658.

Oouchi, Kazuyoshi, Akira T. Noda, Masaki Satoh, Bin Wang, Shang-Ping Xie, Hiroshi G. Takahashi, and Tetsuzo Yasunari (2009). "Asian summer monsoon simulated by a global cloud-system-resolving model: Diurnal to intra-seasonal variability". *Geophysical Research Letters* 36.11. DOI: https://doi.org/10.1029/2009GL038271.

Palmer, Tim and Bjorn Stevens (2019). "The scientific challenge of understanding and estimating climate change". *Proceedings of the National Academy of Sciences* 116.49, pp. 24390–24395. DOI: 10.1073/pnas.1906691116.

Pantillon, F., S. Davolio, E. Avolio, C. Calvo-Sancho, D. S. Carrió, S. Dafis, et al. (2024). "The crucial representation of deep convection for the cyclogenesis of Medicane Ianos". *Weather and Climate Dynamics* 5.3, pp. 1187–1205. DOI: 10.5194/wcd-5-1187-2024.

Parker, Wendy S. (2018). "The Significance of Robust Climate Projections". In: *Climate Modelling: Philosophical and Conceptual Issues*. Ed. by Elisabeth A. Lloyd and Eric

Winsberg. Cham: Springer International Publishing, pp. 273–296. DOI: 10.1007/978-3-319-65058-6_9.

Peinado-Bravo, Angel, Daniel Klocke, and Bjorn Stevens (Aug. 2024). "Horizontal grid spacing convergence of aquaplanet experiments".

Phillips, Norman A. (1956). "The general circulation of the atmosphere: A numerical experiment". *Quarterly Journal of the Royal Meteorological Society* 82.352, pp. 123–164. DOI: https://doi.org/10.1002/qj.49708235202.

Phillips, Tyrone S. and Christopher J. Roy (Sept. 2014). "Richardson Extrapolation-Based Discretization Uncertainty Estimation for Computational Fluid Dynamics". *Journal of Fluids Engineering* 136.12. 121401. ISSN: 0098-2202. DOI: 10.1115/1.4027353.

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

Popp, Max, Nicholas J. Lutsko, and Sandrine Bony (2020). "The Relationship Between Convective Clustering and Mean Tropical Climate in Aquaplanet Simulations". *Journal of Advances in Modeling Earth Systems* 12.8, e2020MS002070. DOI: https://doi.org/10.1029/2020MS002070.

Possolo, Antonio and Hari K. Iyer (Jan. 2017). "Invited Article: Concepts and tools for the evaluation of measurement uncertainty". *Review of Scientific Instruments* 88.1, p. 011301. ISSN: 0034-6748. DOI: 10.1063/1.4974274.

Prein, A. F., A. Gobiet, H. Truhetz, K. Keuler, K. Goergen, C. Teichmann, et al. (Jan. 2016). "Precipitation in the EURO-CORDEX 0.11° and 0.44° simulations: high resolution, high benefits?" *Climate Dynamics* 46.1, pp. 383–412. ISSN: 1432-0894. DOI: 10.1007/s00382-015-2589-y.

Rajedran, Kavirajan, Akio Kitoh, and Jayaraman Srinivasan (2013). "Effect of SST Variation on ITCZ in APE Simulations". *Journal of the Meteorological Society of Japan. Ser. II* 91A, pp. 195–215. DOI: 10.2151/jmsj.2013-A06.

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". *Journal of Advances in Modeling Earth Systems* 11.7, pp. 2070–2088. DOI: https://doi.org/10.1029/2019MS001677.

Richardson, Lewis Fry (2007). *Weather Prediction by Numerical Process*. 2nd ed. Cambridge Mathematical Library. Cambridge University Press.

Rieck, Malte, Louise Nuijens, and Bjorn Stevens (2012). "Marine Boundary Layer Cloud Feedbacks in a Constant Relative Humidity Atmosphere". *Journal of the Atmospheric Sciences* 69.8, pp. 2538–2550. DOI: 10.1175/JAS-D-11-0203.1.

Rio, Catherine, Anthony D. Del Genio, and Frédéric Hourdin (June 2019). "Ongoing Breakthroughs in Convective Parameterization". *Current Climate Change Reports* 5.2, pp. 95–111. ISSN: 2198-6061. DOI: 10.1007/s40641-019-00127-w.

Rios-Berrios, R., B. Medeiros, and G. H. Bryan (2020). "Mean Climate and Tropical Rainfall Variability in Aquaplanet Simulations Using the Model for Prediction Across Scales-Atmosphere". *Journal of Advances in Modeling Earth Systems* 12.10, e2020MS002102. DOI: https://doi.org/10.1029/2020MS002102.

Rios-Berrios, Rosimar, George H. Bryan, Brian Medeiros, Falko Judt, and Wei Wang (2022). "Differences in Tropical Rainfall in Aquaplanet Simulations With Resolved or Parameterized Deep Convection". *Journal of Advances in Modeling Earth Systems* 14.5, e2021MS002902. DOI: https://doi.org/10.1029/2021MS002902.

Roache, Patrick J., Kirti N. Ghia, and Frank M. White (Mar. 1986). "Editorial Policy Statement on the Control of Numerical Accuracy". *Journal of Fluids Engineering* 108.1, pp. 2–2. ISSN: 0098-2202. DOI: 10.1115/1.3242537.

Roemer, Florian E., Stefan A. Buehler, Manfred Brath, Lukas Kluft, and Viju O. John (May 2023). "Direct observation of Earth's spectral long-wave feedback parameter". *Nature Geoscience* 16.5, pp. 416–421. ISSN: 1752-0908. DOI: 10.1038/s41561-023-01175-6.

Sato, Tomonori, Takao Yoshikane, Masaki Satoh, Hiroaki Miura, and Hatsuki Fujinami (2008). "Resolution Dependency of the Diurnal Cycle of Convective Clouds over the Tibetan Plateau in a Mesoscale Model". *Journal of the Meteorological Society of Japan. Ser. II* 86A, pp. 17–31. DOI: 10.2151/jmsj.86A.17.

Satoh, Masaki, Bjorn Stevens, Falko Judt, Marat Khairoutdinov, Shian-Jiann Lin, William M. Putman, and Peter Düben (Sept. 2019). "Global Cloud-Resolving Models". *Current Climate Change Reports* 5.3, pp. 172–184. ISSN: 2198-6061. DOI: 10.1007/s40641-019-00131-0.

Satoh, Masaki, Hirofumi Tomita, Hisashi Yashiro, Yoshiyuki Kajikawa, Yoshiaki Miyamoto, Tsuyoshi Yamaura, et al. (Apr. 2017). "Outcomes and challenges of global high-resolution non-hydrostatic atmospheric simulations using the K computer". *Progress in Earth and Planetary Science* 4.1, p. 13. ISSN: 2197-4284. DOI: 10.1186/s40645-017-0127-8.

Schmidt, H., S. Rast, J. Bao, S.-W. Fang, D. Jimenez-de la Cuesta, P. Keil, et al. (2023). "Effects of vertical grid spacing on the climate simulated in the ICON-Sapphire global storm-resolving model". *EGUsphere* 2023, pp. 1–34. DOI: 10.5194/egusphere-2023-1575.

Schulz, Hauke and Bjorn Stevens (2023). "Evaluating Large-Domain, Hecto-Meter, Large-Eddy Simulations of Trade-Wind Clouds Using EUREC4A Data". *Journal of Advances in Modeling Earth Systems* 15.10, e2023MS003648. ISSN: 1942-2466. DOI: 10.1029/2023MS003648. (Visited on 10/18/2023).

Shaw, Tiffany A. and Aiko Voigt (2016). "Understanding the Links between Subtropical and Extratropical Circulation Responses to Climate Change Using Aquaplanet Model Simulations". *Journal of Climate* 29.18, pp. 6637–6657. DOI: 10.1175/JCLI-D-16-0049.1.

Sherwood, S. C., M. J. Webb, J. D. Annan, K. C. Armour, P. M. Forster, J. C. Hargreaves, et al. (2020). "An Assessment of Earth's Climate Sensitivity Using Multiple Lines of Evidence". *Reviews of Geophysics* 58.4. e2019RG000678 2019RG000678, e2019RG000678. DOI: https://doi.org/10.1029/2019RG000678.

Silvers, Levi G., Kevin A. Reed, and Allison A. Wing (2023). "The Response of the Large-Scale Tropical Circulation to Warming". *Journal of Advances in Modeling Earth Systems* 15.3, e2021MS002966. DOI: https://doi.org/10.1029/2021MS002966.

Skamarock, William C., Joseph B. Klemp, Michael G. Duda, Laura D. Fowler, Sang-Hun Park, and Todd D. Ringler (2012). "A Multiscale Nonhydrostatic Atmospheric Model Using Centroidal Voronoi Tesselations and C-Grid Staggering". *Monthly Weather Review* 140.9, pp. 3090–3105. DOI: 10.1175/MWR-D-11-00215.1.

Smagorinsky, Joseph (1963). "General circulation experiments with the primitive equations: I. The basic experiment". *Monthly weather review* 91.3, pp. 99–164.

Son, Seok-Woo, Seo-Yeon Kim, and Seung-Ki Min (2018). "Widening of the Hadley Cell from Last Glacial Maximum to Future Climate". *Journal of Climate* 31.1, pp. 267–281. DOI: 10.1175/JCLI-D-17-0328.1.

Song, Jinyan, Fengfei Song, Zhe Feng, L. Ruby Leung, Chao Li, and Lixin Wu (2024). "Realistic Precipitation Diurnal Cycle in Global Convection-Permitting Models by Resolv-

ing Mesoscale Convective Systems". *Geophysical Research Letters* 51.13, e2024GL109945. DOI: https://doi.org/10.1029/2024GL109945.

Staniforth, Andrew and John Thuburn (2012). "Horizontal grids for global weather and climate prediction models: a review". *Quarterly Journal of the Royal Meteorological Society* 138.662, pp. 1–26. DOI: https://doi.org/10.1002/qj.958.

Stevens, Bjorn, Andrew S Ackerman, Bruce A Albrecht, Andrew R Brown, Andreas Chlond, Joan Cuxart, et al. (2001). "Simulations of Trade Wind Cumuli under a Strong Inversion". *Journal of the Atmospheric Sciences* 58, p. 22.

Stevens, Bjorn, Claudia Acquistapace, Akio Hansen, Rieke Heinze, Carolin Klinger, Daniel Klocke, Harald Rybka, Wiebke Schubotz, Julia Windmiller, Panagiotis Adamidis, et al. (2020). "The Added Value of Large-eddy and Storm-resolving Models for Simulating Clouds and Precipitation". *Journal of the Meteorological Society of Japan. Ser. II* 98.2, pp. 395–435. DOI: 10.2151/jmsj.2020-021.

Stevens, Bjorn and Sandrine Bony (May 2013a). "What Are Climate Models Missing?" *Science* 340.6136, pp. 1053–1054. ISSN: 1095-9203. DOI: 10.1126/science.1237554.

– (2013b). "What Are Climate Models Missing?" *Science* 340.6136, pp. 1053–1054. URL: https://www.science.org/doi/abs/10.1126/science.1237554.

Stevens, Bjorn, Sandrine Bony, and Mark J. Webb (2012). "Clouds On-Off Klimate Intercomparison Experiment (COOKIE)". In: URL: https://api.semanticscholar.org/CorpusID:18485339.

Stevens, Bjorn, Masaki Satoh, Ludovic Auger, Joachim Biercamp, Christopher S. Bretherton, Xi Chen, et al. (Sept. 2019). "DYAMOND: the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains". *Progress in Earth and Planetary Science* 6.1, p. 61. ISSN: 2197-4284. DOI: 10.1186/s40645-019-0304-z.

Takasuka, Daisuke and Masaki Satoh (2021). "Diversity of the Madden–Julian Oscillation: Initiation Region Modulated by the Interaction between the Intraseasonal and Interannual Variabilities". *Journal of Climate* 34.6, pp. 2297–2318. DOI: 10.1175/JCLI-D-20-0688.1.

Takasuka, Daisuke, Masaki Satoh, Tomoki Miyakawa, Chihiro Kodama, Daniel Klocke, Bjorn Stevens, Pier Luigi Vidale, and Christopher R Terai (May 2024). "A protocol and analysis of year-long simulations of global storm-resolving models and beyond".

Talib, Joshua, Steven J. Woolnough, Nicholas P. Klingaman, and Christopher E. Holloway (2018). "The Role of the Cloud Radiative Effect in the Sensitivity of the Intertropical Convergence Zone to Convective Mixing". *Journal of Climate* 31.17, pp. 6821–6838. DOI: 10.1175/JCLI-D-17-0794.1.

Terai, Christopher R., Peter M. Caldwell, Stephen A. Klein, Qi Tang, and Marcia L. Branstetter (May 2018). "The atmospheric hydrologic cycle in the ACME v0.3 model". *Climate Dynamics* 50.9, pp. 3251–3279. ISSN: 1432-0894. DOI: 10.1007/s00382-017-3803-x.

Tomita, H., H. Miura, S. Iga, T. Nasuno, and M. Satoh (2005). "A global cloud-resolving simulation: Preliminary results from an aqua planet experiment". *Geophysical Research Letters* 32.8. DOI: 10.1029/2005GL022459.

Trenberth, Kevin E., Aiguo Dai, Roy M. Rasmussen, and David B. Parsons (2003). "The Changing Character of Precipitation". *Bulletin of the American Meteorological Society* 84.9, pp. 1205–1218. DOI: 10.1175/BAMS-84-9-1205.

Vallis, Geoffrey K., Pablo Zurita-Gotor, Cameron Cairns, and Joseph Kidston (2015). "Response of the large-scale structure of the atmosphere to global warming". *Quarterly Journal of the Royal Meteorological Society* 141.690, pp. 1479–1501. DOI: https://doi.org/10.1002/qj.2456.

Vergara-Temprado, Jesús, Nikolina Ban, Davide Panosetti, Linda Schlemmer, and Christoph Schär (2020). "Climate Models Permit Convection at Much Coarser Resolutions Than Previously Considered". *Journal of Climate* 33.5, pp. 1915–1933. DOI: 10.1175/JCLI-D-19-0286.1.

Voigt, Aiko, Nicole Albern, Paulo Ceppi, Kevin Grise, Ying Li, and Brian Medeiros (2021). "Clouds, radiation, and atmospheric circulation in the present-day climate and under climate change". *WIREs Climate Change* 12.2, e694. DOI: https://doi.org/10.1002/wcc.694.

Voigt, Aiko and Tiffany A. Shaw (Feb. 2015). "Circulation response to warming shaped by radiative changes of clouds and water vapour". *Nature Geoscience* 8.2, pp. 102–106. ISSN: 1752-0908. DOI: 10.1038/ngeo2345.

Wan, H., M. A. Giorgetta, G. Zängl, M. Restelli, D. Majewski, L. Bonaventura, et al. (2013). "The ICON-1.2 hydrostatic atmospheric dynamical core on triangular grids – Part 1: Formulation and performance of the baseline version". *Geoscientific Model Development* 6.3, pp. 735–763. DOI: 10.5194/gmd-6-735-2013.

Watt-Meyer, Oliver and Dargan M. W. Frierson (2019). "ITCZ Width Controls on Hadley Cell Extent and Eddy-Driven Jet Position and Their Response to Warming". *Journal of Climate* 32.4, pp. 1151–1166. DOI: 10.1175/JCLI-D-18-0434.1.

Webb, M., C. Senior, S. Bony, and J.-J. Morcrette (Nov. 2001). "Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models". *Climate Dynamics* 17.12, pp. 905–922. ISSN: 1432-0894. DOI: 10.1007/s003820100157.

Webb, M. J., T. Andrews, A. Bodas-Salcedo, S. Bony, C. S. Bretherton, R. Chadwick, et al. (2017). "The Cloud Feedback Model Intercomparison Project (CFMIP) contribution to CMIP6". *Geoscientific Model Development* 10.1, pp. 359–384. DOI: 10.5194/gmd-10-359-2017.

Williamson, David L. (2008). "Convergence of aqua-planet simulations with increasing resolution in the Community Atmospheric Model, Version 3". *Tellus A* 60.5, pp. 848–862. DOI: https://doi.org/10.1111/j.1600-0870.2008.00339.x.

Williamson, David L., Michael Blackburn, Kensuke Nakajima, Wataru Ohfuchi, Yoshiyuki O. Takahashi, Yoshi-Yuki Hayashi, et al. (2013). "The Aqua-Planet Experiment (APE): Response to Changed Meridional SST Profile". *Journal of the Meteorological Society of Japan. Ser. II* 91A, pp. 57–89. DOI: 10.2151/jmsj.2013-A03.

Wing, A. A., K. A. Reed, M. Satoh, B. Stevens, S. Bony, and T. Ohno (2018). "Radiative–convective equilibrium model intercomparison project". *Geoscientific Model Development* 11.2, pp. 793–813. DOI: 10.5194/gmd-11-793-2018.

Zängl, Günther, Daniel Reinert, Pilar Ripodas, and Michael Baldauf (2015). "The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core". *Quarterly Journal of the Royal Meteorological Society* 141.687, pp. 563–579. DOI: https://doi.org/10.1002/qj.2378.

Zarzycki, C. M., C. Jablonowski, J. Kent, P. H. Lauritzen, R. Nair, K. A. Reed, et al. (2019). "DCMIP2016: the splitting supercell test case". *Geoscientific Model Development* 12.3, pp. 879–892. DOI: 10.5194/gmd-12-879-2019.

Zelinka, Mark D. and Dennis L. Hartmann (2010). "Why is longwave cloud feedback positive?" *Journal of Geophysical Research: Atmospheres* 115.D16. DOI: https://doi.org/10.1029/2010JD013817.

Zelinka, Mark D., Stephen A. Klein, Yi Qin, and Timothy A. Myers (2022). "Evaluating Climate Models' Cloud Feedbacks Against Expert Judgment". *Journal of Geophysical*

Research: Atmospheres 127.2, e2021JD035198. DOI: https://doi.org/10.1029/2021JD035198.

Zelinka, Mark D., Timothy A. Myers, Daniel T. McCoy, Stephen Po-Chedley, Peter M. Caldwell, Paulo Ceppi, Stephen A. Klein, and Karl E. Taylor (2020). "Causes of Higher Climate Sensitivity in CMIP6 Models". *Geophysical Research Letters* 47.1, e2019GL085782. DOI: https://doi.org/10.1029/2019GL085782.

VERSICHERUNG AN EIDES STATT

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Sofern im Zuge der Erstellung der vorliegenden Dissertationsschrift generative Künstliche Intelligenz (gKI) basierte elektronische Hilfsmittel verwendet wurden, versichere ich, dass meine eigene Leistung im Vordergrund stand und dass eine vollständige Dokumentation aller verwendeten Hilfsmittel gemäß der Guten wissenschaftlichen Praxis vorliegt. Ich trage die Verantwortung für eventuell durch die gKI generierte fehlerhafte oder verzerrte Inhalte, fehlerhafte Referenzen, Verstöße gegen das Datenschutz- und Urheberrecht oder Plagiate.

Hamburg, 10. December 2024

Angel Peinado Bravo

I hereby declare and affirm that this doctoral dissertation is my own work and that I have not used any aids and sources other than those indicated.
If electronic resources based on generative artificial intelligence (gAI) were used in the course of writing this dissertation, I confirm that my own work was the main and value-adding contribution and that complete documentation of all resources used is available in accordance with good scientific practice. I am responsible for any erroneous or distorted content, incorrect references, violations of data protection and copyright law or plagiarism that may have been generated by the gAI.

Hamburg, 10. Dezember 2024

Angel Peinado Bravo