



# Early Development and Tuning of a Global Coupled Cloud Resolving Model, and its Fast Response to Increasing CO<sub>2</sub>

THORSTEN MAURITSEN

RENE REDLER

MONIKA ESCH

BJORN STEVENS

CATHY HOHENEGGER

DANIEL KLOCKE

RENATE BROKOPF

HELMUTH HAAK

LEONIDAS LINARDAKIS

NIKLAS RÖBER

REINER SCHNUR

\*Author affiliations can be found in the back matter of this article

ORIGINAL RESEARCH  
PAPER



STOCKHOLM  
UNIVERSITY PRESS

## ABSTRACT

Since the dawn of functioning numerical dynamical atmosphere- and ocean models, their resolution has steadily increased, fed by an exponential growth in computational capabilities. However, because resolution of models is at all times limited by computational power a number of mostly small-scale or micro-scale processes have to be parameterised. Particularly those of atmospheric moist convection and ocean eddies are problematic when scientists seek to interpret output from model experiments. Here we present the first coupled ocean-atmosphere model experiments with sufficient resolution to dispose of moist convection and ocean eddy parameterisations. We describe the early development and discuss the challenges associated with conducting the simulations with a focus on tuning the global mean radiation balance in order to limit drifts. A four-month experiment with quadrupled CO<sub>2</sub> is then compared with a ten-member ensemble of low-resolution simulations using MPI-ESM1.2-LR. We find broad similarities of the response, albeit with a more diversified spatial pattern with both stronger and weaker regional warming, as well as a sharpening of precipitation in the inter tropical convergence zone. These early results demonstrate that it is already now possible to learn from such coupled model experiments, even if short by nature.

## CORRESPONDING AUTHOR:

**Thorsten Mauritsen**

Meteorologiska Institutionen  
vid Stockholms Universitet  
(MISU), Stockholm, Sweden

[thorsten.mauritsen@misu.su.se](mailto:thorsten.mauritsen@misu.su.se)

## KEYWORDS:

Climate model; cloud resolving

## TO CITE THIS ARTICLE:

Mauritsen, T, Redler, R, Esch, M, Stevens, B, Hohenegger, C, Klocke, D, Brokopf, R, Haak, H, Linardakis, L, Röber, N and Schnur, R. 2022. Early Development and Tuning of a Global Coupled Cloud Resolving Model, and its Fast Response to Increasing CO<sub>2</sub>. *Tellus A: Dynamic Meteorology and Oceanography*, 74(2022): 346–363. DOI: <https://doi.org/10.16993/tellusa.54>

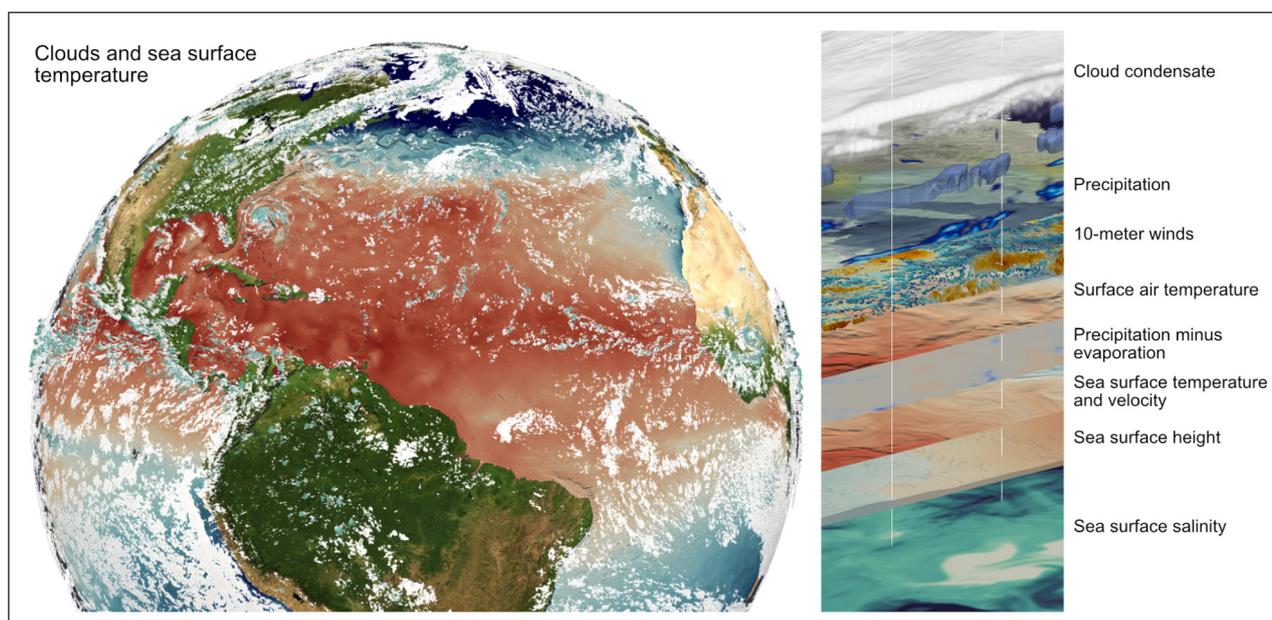
## 1 INTRODUCTION

Modern coupled ocean-atmosphere climate modeling has its roots in the idea that one can simulate the motion of the atmosphere and oceans using the laws of physics. This idea dates back more than a century when Bjerknes (1904) first proposed weather forecasting as an initial value problem. His idea was quickly followed up by Richardson (1922) in his seminal attempt to calculate a short weather forecast by hand. With the advent of computers such simulations became applicable to both weather forecasting and climate modeling in the 1950's and 1960's (Charney et al., 1950; Bolin, 1955; Phillips, 1956; Manabe et al., 1965; Manabe and Bryan, 1969). A limitation of climate modeling is that important small scale motions, not resolved by the computational grid, must be parameterised which is a leading source of uncertainty and a limitation to our ability to understand the results (Hohenegger and Stevens, 2018; Retsch et al., 2019; Hohenegger et al., 2020; Uribe et al., 2021). In this paper we describe the development of a coupled climate model with sufficient resolution to represent atmospheric moist convection, gravity wave drag and ocean eddies and so can dispose of their parameterisations (Figure 1).

Climate modeling, as an activity on its own, has not come far over the past five decades in answering basic questions such as how much warmer the planet might be at the end of this century (Zelinka et al., 2020; Flynn and Mauritsen, 2020), and likewise progress on representing regional information on societally important quantities such as precipitation change has been close to non-existent (Shepherd, 2014; Fiedler et al., 2020). A common approach in the community is to further increase the complexity of models and to elaborate their

parameterisations (Washington et al., 2008). Although such continual refinement may help to better fit aspects of the observed climate, and these types of models will undoubtedly remain useful tools for decades to come (Balaji et al., 2022), the idea that major breakthroughs are to be expected has been challenged (Palmer and Stevens, 2019). Another recent idea is to replace the model parameterisations with machine learning algorithms (Schneider et al., 2017), though this approach has still to be demonstrated in actual experiments. Furthermore, testing and interpreting results from such models on the climate change problem may prove challenging too.

A more transparent approach, the one which we pursue here, is to dramatically increase the model grid resolution to the point where parameterisations can be reduced in number, and for those that inevitably remain, more accurate versions can be chosen. There is vast experience that this approach can be fruitful from studies of limited area simulations, or stand alone atmosphere or ocean models (e.g. Deardorff, 1970; Klemp and Wilhelmson, 1978; Smith et al., 2000; Tomita et al., 2005; Heinze et al., 2017; Stevens et al., 2019; Retsch et al., 2019; Uribe et al., 2021). In particular horizontal ocean eddies, moist convection and various forms of gravity wave drag are for the most part parameterised in contemporary climate models, but at kilometer scale resolutions the effect of these processes can largely be represented by resolved motion. That is not to say that these processes are by any means fully resolved when using such a grid (e.g. Radtke et al. 2021), rather the idea is that a distorted representation of a physical process, one based on the solution of the basic equations that govern it, is often easier to interpret and understand.

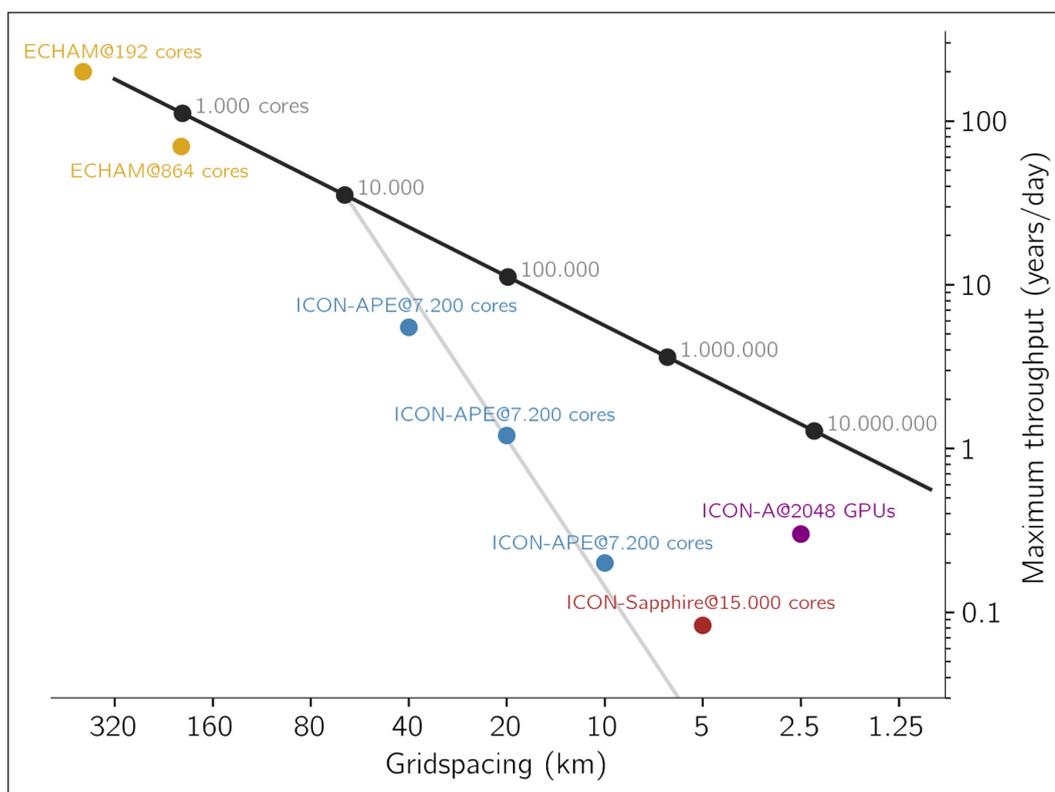


**Figure 1** Snapshot of sea surface temperature and clouds in the coupled ICON-Sapphire experiments. Left is the North Atlantic region with sea surface temperatures displayed in colors from warm (red) to cold (blue) and a three-dimensional volume rendering of clouds. To the right is a layered display of various variables in the model. Note that several of these are defined at the surface.

We will argue that the time to develop global cloud resolving climate models is now. Experimentation with such models is by and large limited by the number of years that can be simulated on computers available to scientists in a given amount of real time, also referred to as throughput, or temporal compression, typically measured in days per day or years per day. Today, practical implementations divide the problem into smaller pieces that are then calculated in parallel on individual compute cores which then exchange data to solve the global problem. The maximum throughput that can be attained at a given resolution if the number of cores is infinite is limited foremost by time-step length, the speed of individual cores, and communication (Amdahl, 1967). For the ICON model and a recently retired computer architecture we estimate the theoretical maximum throughput and also display a series of real world examples (Figure 2). The maximum throughput limit can be increased, although not dramatically. The experiments with ICON discussed in this paper were done at a throughput of about 17–30 days per day

depending on configuration, although in practice only a fraction of this because the simulations had to queue on the computer, and as such these experiments were at the limit of what was feasible for us to do up to now. Nevertheless, with the rapidly decreasing cost of compute resources (Moore, 1965), such simulations are going to become more common in the coming years.

Also, technical innovation, such as the use of graphics processing units (GPU) can accelerate progress (Yashiro et al., 2016; Fuhrer et al., 2018); a current example from ICON is given in Figure 2. These GPU-based simulations will have a lower maximum throughput due to their inherent parallelism and low per-core performance. The example given is close to the maximum throughput at 2.5 km resolution, about a factor 4–5 below the theoretical maximum throughput with CPUs. To achieve the corresponding throughput to these 2048 GPUs, however, we estimate we would need about 1 million CPU cores. As such GPUs present an advantage over CPUs on these large problems that are limited by the available memory and computing power.



**Figure 2** The maximum throughput in years per day as a function of horizontal grid spacing. Thick black line is the estimated maximum with recent technology, here the Intel Xeon E5-2695V4 Broadwell processor based Mistral supercomputer built 2015 at the Deutsches Klimarechenzentrum (DKRZ). The slope of this line is determined by the time step length at a given resolution. Dots and numbers along that line indicates the approximate number of compute cores needed to reach that level of performance as extrapolated from the low-resolution experiments. Grey line shows the approximate performance for a given number of cores at increasing resolutions, as can be compared with the blue symbols: Yellow and blue symbols are for ECHAM6.3 and ICON atmosphere-only experiments carried out in 2017 without output, whereby ICON-A and ECHAM6.3 experiments were with continents and ICON-APE is idealised aquaplanet experiments. The brown symbol is for the coupled ocean-atmosphere ICON-Sapphire configuration as used in this study with twice as many levels as the aquaplanet experiments and asynchronous output. This coupled model run contains many optimisations over the earlier ICON-APE experiments, but also dedicate compute cores to an ocean and is hampered by some load unbalancing. Purple symbol is an atmosphere only experiment that has been ported to using Nvidia A100 Graphics Processing Units (GPUs) on the Jülich JUWELS Booster supercomputer.

Although experimentation with globally coupled cloud-resolving models in the 2020s will for the most part be limited to decades, and one must dismiss the idea of eliminating long term climate drifts through millennia long spin-ups such as is commonly applied to current climate models (Mauritsen et al., 2012; Hourdin et al., 2017), there is still a wealth of interesting experiments that can be conducted and phenomena that can be studied which were not feasible before. What is more, due to the maximum throughput limitation, these simulations will probably not be able to run at more than a few years per day within the foreseeable future, such that their scope to study timescales longer than a few hundred years is inherently limited, i.e. that which can be computed in a year. Hence, we argue, the time to develop, apply and exploit these new tools is now.

In the following we will share our experience with developing ICON into a cloud- and ocean eddy resolving model that resulted in multiple year-long simulations, and thereafter we shall investigate the models surface temperature and precipitation response in the first four months following a quadrupled atmospheric CO<sub>2</sub> concentration, demonstrating how such models can already now be used to gain understanding.

## 2 MODEL DEVELOPMENTS

The purpose of the here described project, launched in the winter 2017/18, was to demonstrate that coupled simulations with sufficient resolution to explicitly represent moist deep convection and ocean eddies are now both feasible and useful. As such the purpose was not in the first place to achieve fidelity with observations in all respects, but first and foremost to show that already now it is possible to experiment with a coupled ocean-atmosphere model at this level of resolution. To this end, we decided to aim for running an annual cycle simulation during the project, and in addition it was decided to explore the model response to increasing CO<sub>2</sub>. ICON is developed in a few different configurations, for instance to support numerical weather prediction (ICON-NWP), or as a traditional CMIP class model (ICON-ESM). The model version to be developed here was named ICON-Sapphire, with reference to the gems blue colour, signifying the focus on fine-scale resolution. In this publication we focus on describing the process of the early developments, whereas a complete description of the ICON-Sapphire modeling system and its many other capabilities, as they developed out of this and other projects, is given in a companion paper (Hohenegger et al., 2022).

### 2.1 STRATEGY

To achieve this goal a pragmatic approach was taken – something we came to refer to as “the workbench” – whereby existing model components were first brought

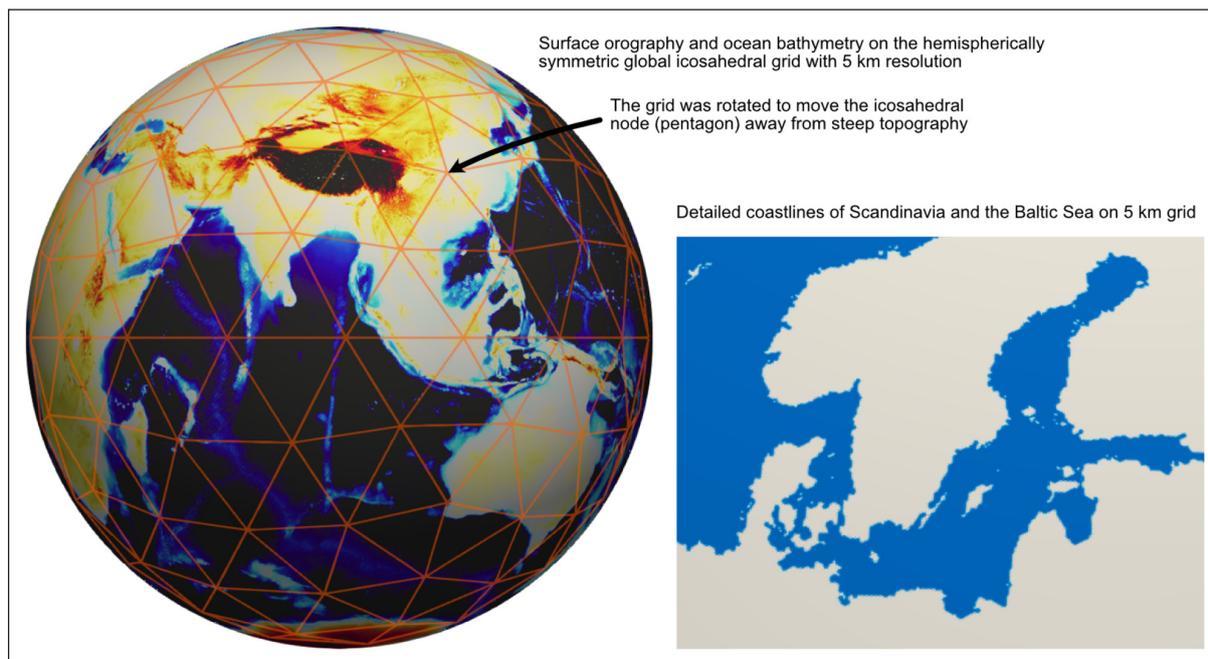
together to a setup that could actually run, even if only for a few time steps before the computation would fail. Thereafter, problems could be identified and amended as they would become apparent. New developments could then be tested in an already working coupled setup. This way, the overall progress of the project did not hinge on a single development. Insofar that it was possible, the latest setup of the model was kept running on the computer at all times to gather information about both technical issues, computational performance and physical biases.

The chosen development strategy was computationally intensive. The coupled simulation required a minimum of 15.000 computational cores of the total 100.000 cores available on the DKRZ Mistral supercomputer in order to fit in the systems memory. This represented a tremendous institutional investment as the Max Planck Institute has access to about half the computer, and so the development had to be weighed against displaced ongoing research also taking place on the same allocation. To hedge against the high development cost, both uncoupled atmosphere-only (5 km) and lower resolution coupled (160 km) setups were used extensively for testing. Although these setups could be used for a number of purposes, some of the problems we faced would only show up in the high resolution coupled setup, and so also gaining experience running this more expensive setup was indispensable to the development.

### 2.2 MATCHING OCEAN AND ATMOSPHERE GRIDS AND THEIR COUPLING

In the spirit of keeping things simple, it was decided to use matching grids in the atmosphere and ocean. This approach eliminates the need to interpolate during coupling and simplifies the handling of coastlines. The grid of ICON is based on the icosahedron (Satoh, 2014), which consists of 20 triangles projected onto the sphere (Figure 3). The sides of the triangles are bisected to form smaller triangles to achieve the desired resolution. To save system memory the land grid points in the ocean model were excluded.

Several measures were taken to make the grid more uniform and improve the dynamics of the model. The triangles near the corners of the original icosahedron form a pentagon, rather than a hexagon (Figure 3), and their edges are the shortest by default. To reduce this grid distortion of the grid a spring dynamics optimisation is applied (Tomita et al., 2001, 2002), leading to reduced numerical errors (Weller et al., 2009; Korn and Linardakis, 2018). Furthermore, out of experience it can cause numerical instabilities, resulting in model crashes, if these points are placed on steep orography. Therefore it was decided to rotate the grid by 37 degrees to move such a point out of the Himalayas (Figure 3). Finally, the grids were symmetrised with respect to the equator (Heikes and Randall, 1995a, b), which has been found to improve the modelled ocean dynamics (Korn and Linardakis, 2018)



**Figure 3** Illustrations of the global hemispherically symmetric 5 km resolution grid used here. Left is a global view of the ocean bathymetry and land orography. Also shown as red lines is the first refinement of the basic icosahedral grid. Note how this version of the grid is stretched to be symmetric about the equator in order to minimise imprints on tropical ocean dynamics, and also how the nodes of the original icosahedron that are surrounded by a pentagon are shifted away from the steep orography of the Himalayas. On the right is shown a zoom of the details of coastlines around the Baltic Sea.

and in preliminary tests with an aquaplanet setup to reduce the frequency of high horizontal Courant-Friedrich-Lewy (CFL) number events in the atmosphere, thereby permitting a slight increase in the time step length.

The ocean and atmosphere model components run concurrently and perform a parallel data exchange. The number of compute processes for the atmosphere and ocean are determined independently, and as a consequence we do not have a one-to-one relation between the ocean and atmosphere processes. To handle the exchange at the ocean-atmosphere interface we use the YAC coupling library (Hanke et al., 2016; Hanke and Redler, 2019). The interpolation capabilities of YAC are not used here, as due to the matching grids a simple nearest-neighbor search to handle the repartitioning is performed during the initialisation. This search takes roughly 5 seconds. During the run, the coupling routines are called at every model time step, whereby surface exchange data are collected. At user-defined coupling events – here every 900 seconds – the data is averaged in time and sent to the respective receiving processes in the atmosphere and ocean. The resulting coupling performance is satisfactory and scales well.

In practice, the setup led to a substantial load imbalance. This happened because at a given resolution the ocean can take longer time steps (120s) than the atmosphere (30–45s), but the memory usage is similar: The ocean would fit in memory with 120 compute nodes, each containing 36 compute cores and 64 GB memory on the Mistral supercomputer; for the atmosphere this memory limitation was 150 nodes. The memory bottleneck of the

ocean was related to the output mechanism used by the ocean, and this issue has later been addressed. In practice the size and total load on the machine meant we would not request more than 300 nodes for the atmosphere. Consequently, the ocean was waiting for the atmosphere at each coupling event. With a larger machine the load imbalance could be easily eliminated since the ICON atmosphere is far from scaled out at 5 km resolution (Figure 2, ICON-Sapphire is well below the black line).

### 2.3 EXPERIMENT OVERVIEW

The development of the physical model could in retrospect be described as having happened along major development steps, which were characterised by being run longer, complemented by shorter intermediate tests. As described above the strategy was to start with something that works, here the recently developed ICON-ESM (Giorgetta et al., 2018; Jungclaus et al. 2021), and then incrementally move towards the goal of a coupled model with parameterisations that are suitable for cloud resolving simulations at kilometre or finer scales.

In defining our goal, we took inspiration from both large-eddy simulations (LES) and cloud resolving models (CRMs), which are both widely used on limited area domains, and we decided to aim at having an advanced cloud micro physics parameterisation (Baldauf et al., 2011), combined with a three-dimensional turbulence mixing scheme (Dipankar et al., 2015). The latter was not introduced from the start since it was still in development, see below. Likewise, the partial cloud fraction scheme Sundqvist et al. (1989) was replaced with a simple binary

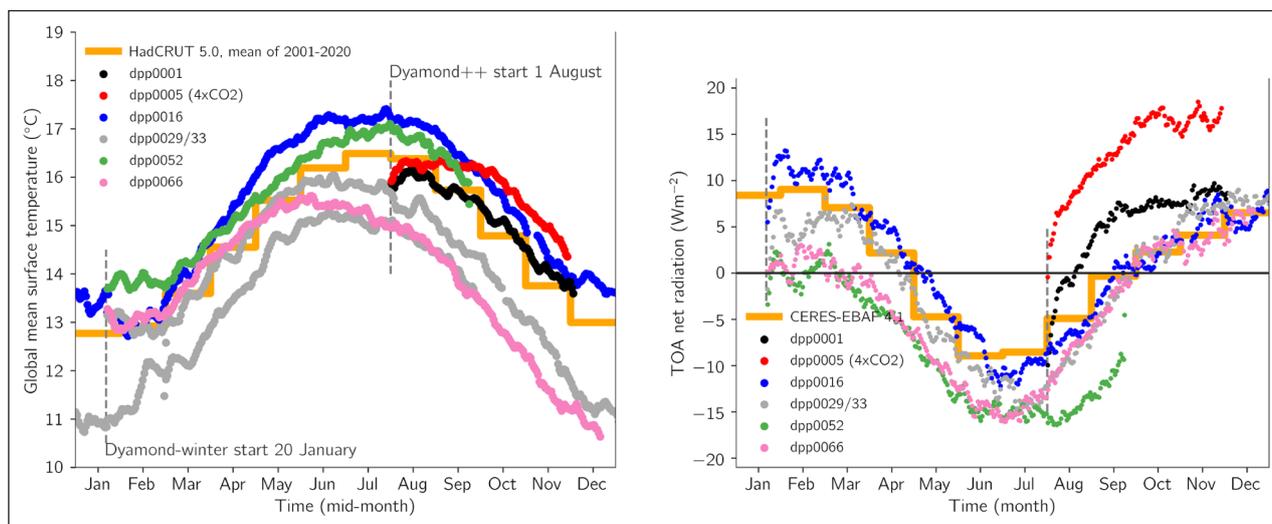
cloud fraction. The convection (Tiedtke, 1989) and gravity wave drag (Lott, 1999) parameterisations were turned off because these processes are assumed to be resolved by the equations of motion. The development could be viewed as having had three phases:

1. Early experiments focused on technical coupling and computational performance
2. Replacement of cloud microphysics (dpp0001/5, dpp0016)
3. Replacement of turbulent mixing scheme (dpp0029/33, dpp0052, dpp0066)

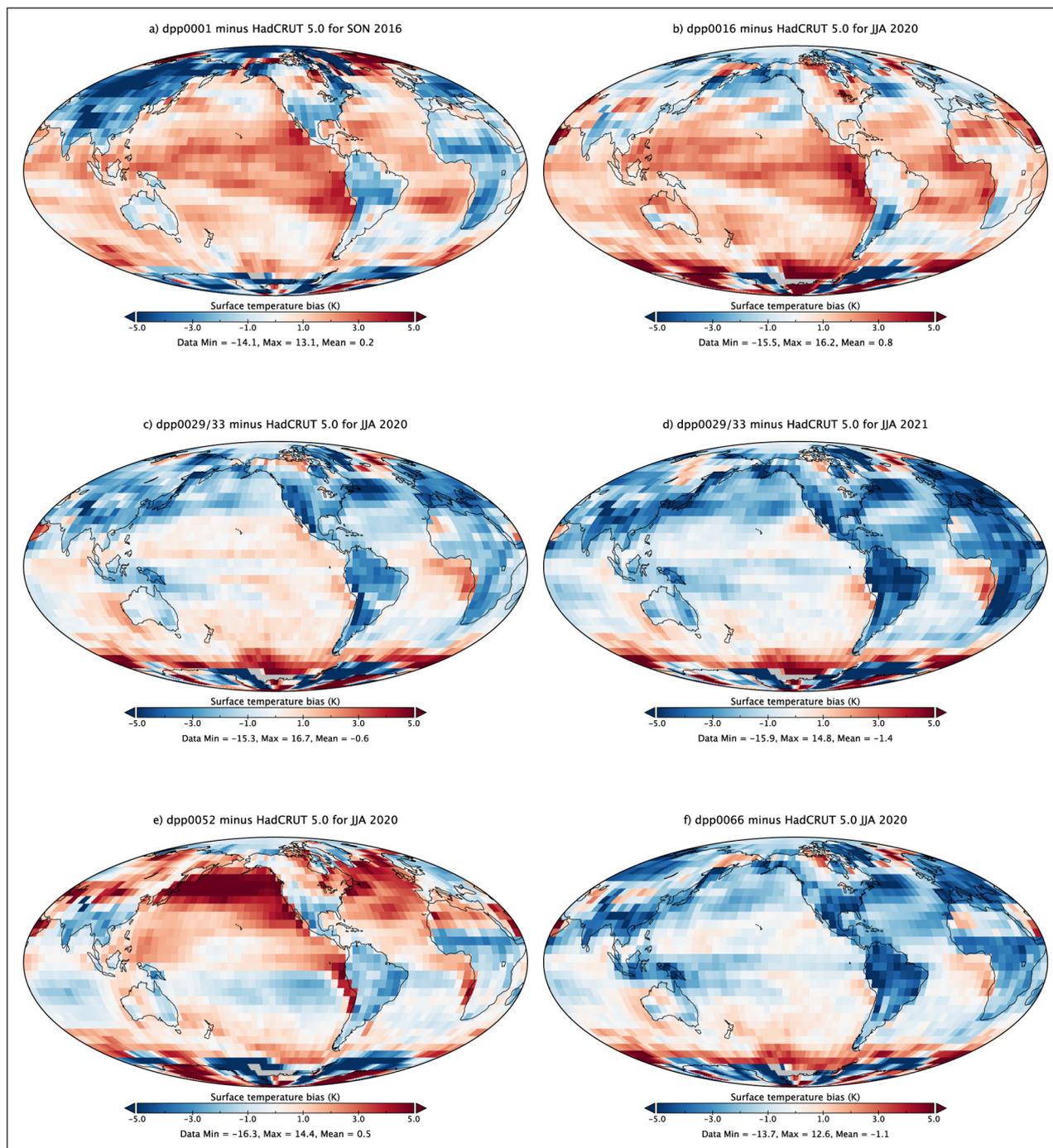
Here the experiment names are “dpp” after DYAMOND++, named after the DYNAMICS of the Atmospheric general circulation Modeled On Non-hydrostatic Domains (DYAMOND) project (Stevens et al., 2019), ‘++’ referring to this being coupled to the oceans rather than using prescribed sea surface temperatures and sea ice distributions, followed by a number in the series of experiments. Each experiment is described in more detail in Table 1. An overview of the temperature and radiation imbalance evolution in the experiments is displayed in Figure 4, and maps of surface temperature biases in Figure 5.

EXPERIMENT ID	DESCRIPTION
dpp0001:	The first working model was based on existing components from ICON-ESM, but with advanced cloud micro physics, convection and gravity wave drag parameterisations turned off, and using a binary cloud fraction scheme. The model was still using the total turbulent energy mixing scheme (Pithan et al., 2015). The experiment was started 1 August 2016 and ran with multiple numerical crashes starting in November until 19 December.
dpp0005:	In the companion experiment to dpp0001 the atmospheric CO <sub>2</sub> was quadrupled (see Section 3).
dpp0016:	To reduce the model warming drift the average ocean surface albedo was raised from 7 to 12 percent (Section 2.5). Also various technical improvements were made to reduce the numerical instabilities. This permitted us to raise the model top from 30 to 75 km. The experiment was started on 20 January 2020, and ran for one year.
dpp0029/33:	In the third step we replaced the total turbulent energy mixing scheme with the Smagorinsky three dimensional turbulence scheme (Dipankar et al. 2015). Furthermore we improved the coupling between winds and ocean currents. The run has substantially more clouds, so the ocean albedo was again reset to its default. In addition the cloud inhomogeneity factor was reduced from 1 to 0.66 to reduce solar reflection. The dpp0029 simulation was started on 20 January 2020 and after reaching one year it was extended as dpp0033 by another 9 months using more vertical resolution in the ocean.
dpp0052:	In this update a programming error in the surface sensible heat flux calculation was removed and a new ocean vertical coordinate whereby a bug in the ocean momentum forcing was introduced. The tuning parameters were kept the same as in dpp0029/33.
dpp0066:	In this update the new ocean vertical coordinate, and the associated bug, introduced in dpp0052 were removed again in order to separate its effect from other changes. The tuning parameters were again kept the same as in dpp0029/33.

**Table 1** Overview of experiments conducted during the project.



**Figure 4** Evolution of temperature and radiation balance with time in the coupled simulations. Left panel shows daily global mean temperature in simulations compared with the average annual cycle from HadCRUT 5.0 reconstruction averaged over the 2001–2020 period (Morice et al. 2021). Starting dates of the experiments are marked with vertical dashed lines. The right panel shows the top of atmosphere radiation balance compared to monthly mean observed radiation balance from CERES-EBAF edition 4.1 averaged over the 2001–2020 period (Loeb et al. 2018). Note that the right panel only shows the first year of dpp0029/33, and also that dpp0005 has quadrupled CO<sub>2</sub> and so is therefore not expected to match observations.



**Figure 5** Surface temperature biases for June to August relative to HadCRUT 5.0 (Morice et al. 2021) averaged over the years 2001–2020. Note that panel **a**) shows September to November (SON), whereas the other panels are June to August (JJA). Panels **c**) and **d**) shows the bias in the first and second year of the dpp0029/33 experiment.

## 2.4 MODEL INITIALISATION

In the first experiments with the new model setup, we frequently experienced numerical instabilities in the early stages of the simulation caused by large gravity waves in the ocean. The ocean was initialised from a climatology which did not have a dynamically balanced surface circulation. Therefore we undertook a 10 year ocean-only simulation at 10 km resolution with initial conditions from the 30 km ORAS5 ocean reanalysis and surface boundary conditions from the ERA-5 atmosphere reanalysis. The resulting state of the ocean was then interpolated to 5

km resolution, and the spin up continued for another 10 years. Although costly, the method resulted in an initial state with balanced ocean eddies, and furthermore we experienced fewer model crashes after the ocean was coupled to the atmosphere.

The atmosphere was initialised from an ECMWF operational analysis on the initial time and day of the experiment. For dpp0001 this was 00 UTC on 1 August 2016 as in Stevens et al. (2019), whereas in the other experiments a starting date of 20 January 2020 was chosen to match the start date of the EUREC4A field

campaign (Bony et al., 2017; Stevens et al., 2021). The initial data for land soil moisture and temperature of the 5 soil layers as well as snow cover are the same as used in the MPI-ESM1.2 model and were simply interpolated to the ICON grid. The procedure was refined in simulations subsequent to dpp0029/33 to instead use soil moisture, soil temperature and snow fields from the same ECMWF analysis with which the atmospheric state is initialised. It is not obvious which approach is better, though, since in both cases these are modelled fields from modeling systems likely to exhibit different snow and soil moisture climates to that of ICON-Sapphire. Thus a drift in these fields during the first years of simulation is inevitable.

## 2.5 TUNING EXPERIENCE, DRIFTS AND BIASES

Drifts in the global mean temperature of coupled climate models is commonly controlled by tuning the radiation balance followed by long spin-ups of typically several thousands of years (Hourdin et al., 2017). If the radiation balance is not tuned the surface temperature will drift away from the observed, thereby making it more challenging to exploit the model for scientific purposes. Tuning is typically done by adjusting various model parameters, mostly pertaining to cloud processes which tend to be effective in controlling the radiation balance. Furthermore, this tuning can also be used to compensate energy leakages in climate models, sometimes on the order of  $5 \text{ Wm}^{-2}$ , by maintaining a correspondingly compensating top of atmosphere radiation imbalance (Mauritsen et al., 2012). There are, however, particular issues to consider when minimising drift in a global cloud resolving model. Here we provide our experience, in the hope that other groups pursuing global coupled cloud resolving models may benefit.

### 2.5.1 Tuning parameters

The parameters used in climate model tuning are usually considered uncertain due to issues with representing small scale cloud processes at grid spacing of the order of hundred kilometers, and thereby justified as tunable parameters. At kilometer scale resolutions, however, these uncertainties are substantially reduced such that one might question the justification. Furthermore, because we turn off the convection scheme in order to instead explicitly resolve convection, a number of parameters that are typically used for tuning are consequently lacking. With fewer and less uncertain parameters at disposal, we may have to resort to using parameters outside their respective ranges of uncertainty, or parameters not usually used for tuning in order to compensate for model structural errors. As an example we used ocean surface albedo to tune dpp0016 colder relative to the warm-biased dpp0001. Although such measures are not justified by an imperfect knowledge of the ocean surface albedo, it may still be justified to do such tuning in order to limit model drift, here caused by a lack of low-level clouds.

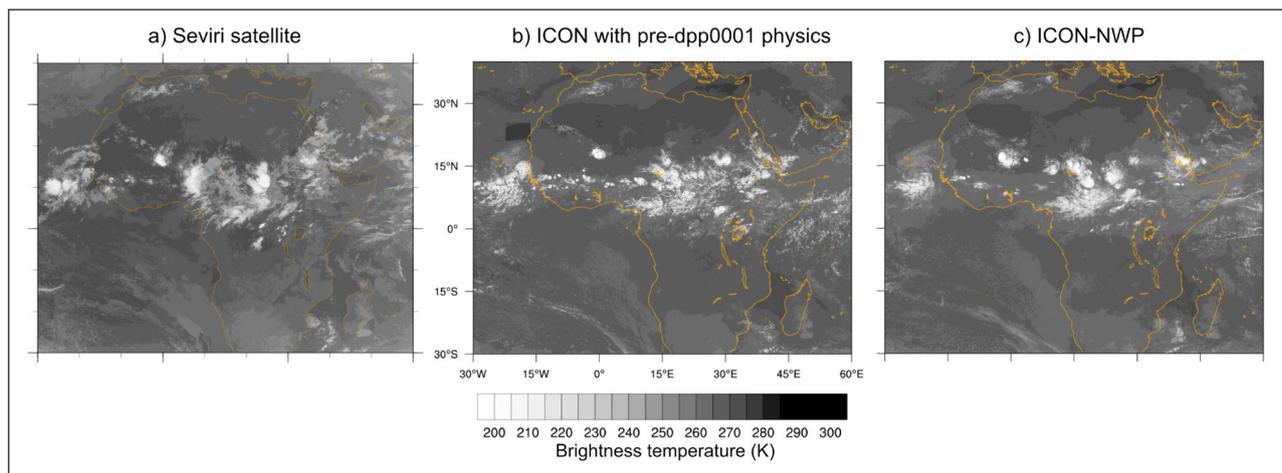
Tuning parameters may not always work the way they used to do at lower resolutions. We encountered such an example during the pre-dpp0001 phase with the relative humidity based fractional cloud cover scheme. In this scheme a critical relative humidity profile, usually 70–90 percent, determines at which large-scale relative humidity clouds start to form; a lower value means more clouds and usually has a cooling effect. We initially set this to 100 percent in our experiments to yield an all-or-nothing scheme, but when we were faced with too few low level clouds in early versions (Phase 1, Section 2.3), we tried to lower the value. To our surprise instead this led to even fewer clouds. Since this route was anyway not aligned with our long term vision we did not investigate the cause further. We speculate, however, that lowering the threshold for cloud formation could cause resolved convection to trigger more easily, leading to a drying of the atmosphere.

### 2.5.2 Informative short runs

To gain experience with the tuning parameters we found it useful to analyse short initialised atmosphere only runs. Usually, when tuning the contemporary climate model MPI-ESM1.2 we run the model in atmosphere only configuration for years or decades to get a good estimate of the parameters effects while averaging over internal variability. An alternative is to use short initialised runs, wherein weather events are nearly the same (Williams et al., 2013; Wan et al., 2014). These runs need to be shorter than the weather prediction limit of about two weeks (Lorenz, 1969), so we decided to use 5-day simulations which ensures similar synoptic scale weather.

We primarily used such short runs to determine how strong the effect of a certain parameter change is on the global mean radiation balance. Since the parameters we used for tuning were either related to cloud processes, or the ocean surface albedo, which have either fast or instantaneous effects our experience was that we were able to obtain a reasonable estimate of their longer term effects based on 5-day simulations.

However, the short runs were also useful in a more qualitative sense. In Figure 6 we compare a satellite image of the longwave brightness temperature, estimated by combining multiple channels, with that which the satellite would have seen in two versions of the ICON model. First, we note a striking similarity: convective clouds are in roughly the right places. But both model versions exhibit smaller convective clouds with less of the thin anvil seen as a light gray veil over central Africa in the satellite image. This was more pronounced in the ICON version shown in panel b) which still used the diagnostic cloud micro physics. The comparison motivated the move to the prognostic scheme used in the numerical weather prediction version, ICON-NWP (panel c). We also noted the box-shaped anomaly near Morocco which was caused by an error in the surface boundary conditions.



**Figure 6** Comparison of infrared brightness temperature observed from the Seviri weather satellite (a), with 12-hour initialised runs from an early version of ICON-Sapphire with diagnostic cloud micro physics (b), and with the numerical weather prediction (NWP) physics package that includes prognostic cloud micro physics used by the German Weather Service (DWD) (c).

### 2.5.3 Controlling drift

Contemporary climate models are typically spun up to reach a stationary state before experimentation starts, however, it is not practical to make such long spin up runs with coupled cloud resolving models. Even if eventually it will no longer be nearly as computationally expensive to run these models, they are still going to be inherently slow, and so it will take years of real time to perform a spin up anywhere nearly long enough to eliminate drifts. Consequently, in addition to limit drifts caused by a biased radiation balance, the development should aim at also reducing initial model drifts which is perhaps a more challenging task.

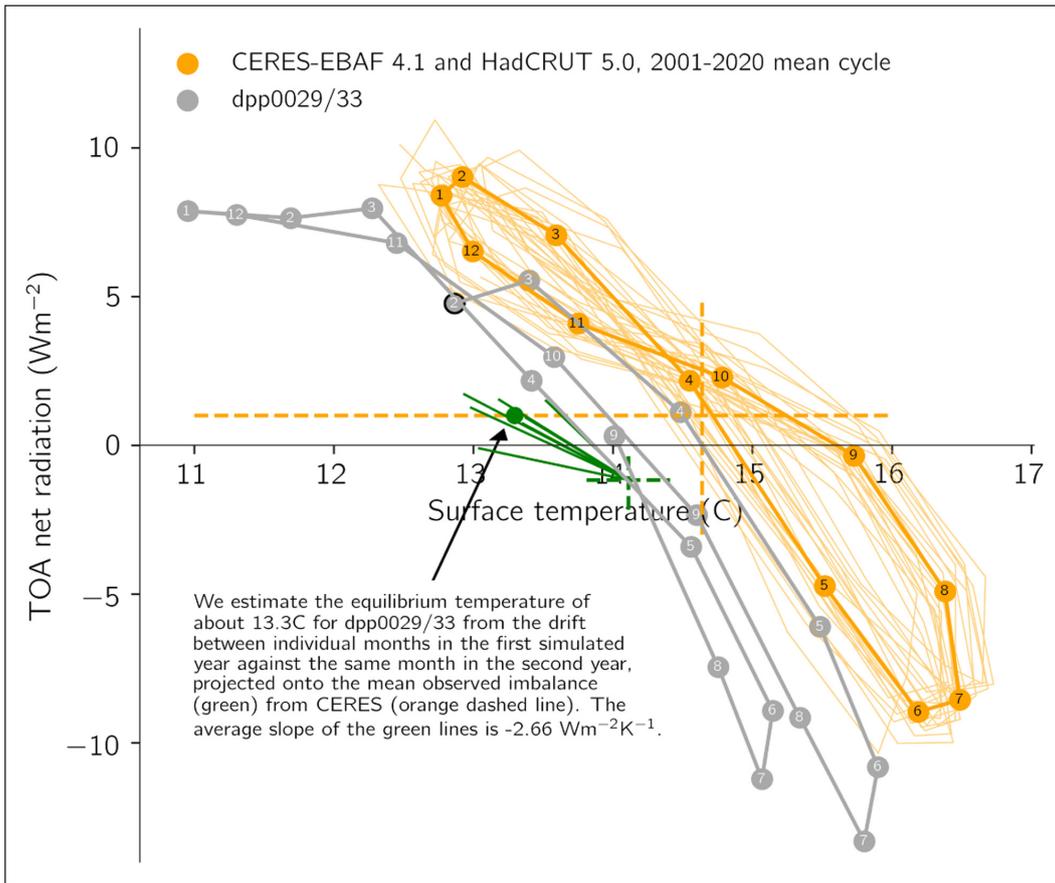
If we assume that the model conserves energy, then we can combine the modelled global mean temperature and radiation balance (Figure 4) to estimate how it will drift. For instance, dpp0001 absorbs more radiation than observed, and consequently it will warm up with time, whereas dpp0029/33 cools due to negative radiation balances. The tuned radiation balance of dpp0016 meant that it did not drift much. An interesting counter example is dpp0052 which has a negatively biased radiation balance, yet is warmer than observed; a case we shall discuss further below.

Estimating at which global mean temperature a model will drift towards is difficult based on short experiments. Here focus is often on the evolution of daily or monthly means (e.g. Figure 4), but biases and drifts in these may result also from an erroneous representation of the annual cycle. From the longer dpp0029/33 experiment we can investigate this in more detail by plotting monthly mean radiation balance against surface temperature (Figure 7). First we notice the observed evolution is shaped as an eight, presumably as surface temperature lags behind the radiation balance due to the heat capacity. The dpp0029/33 run, however, starts at lower radiation balance and hence drifts to colder temperatures, exhibiting a less obvious eight-shaped loop. Comparing

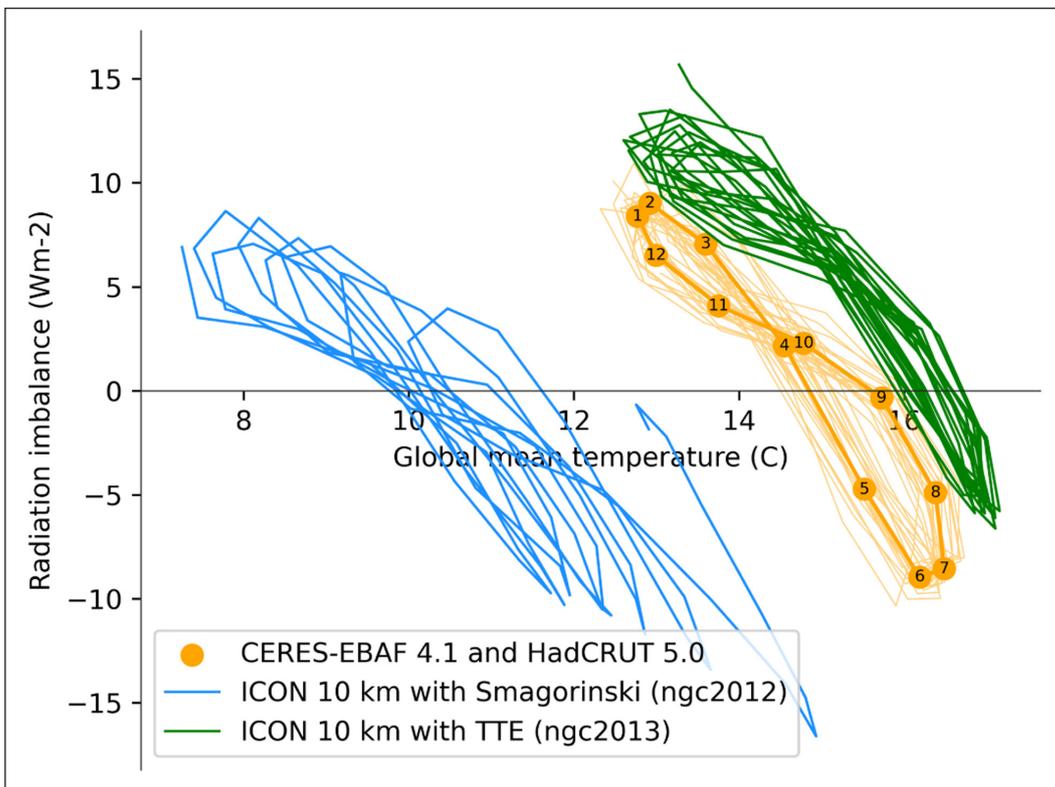
the same months between the first and second year we see that the model drifts on a negative slope, consistent with an overall negative feedback (Mauritsen et al., 2012) which will eventually bring the model into balance at a lower temperature. Based on these slopes we estimate the model will reach such balance for present day boundary conditions approximately 1.5 degrees below observed temperatures.

However, the ICON model versions used here did not conserve energy to varying extents. When there is an artificial energy leakage a model will reach equilibrium when the radiation balance equals the leakage (Mauritsen et al., 2012). Incidentally, we conducted an illustrative pair of 10 year long runs, with 10 km resolution, using the two different turbulence mixing schemes (Figure 8): Smagorinski (ngc2012) and total turbulent energy (ngc2013). Using the latter results in fewer clouds, and hence a substantially more positive radiation balance. The ngc2013 run is practically stable with temperatures close to the observed, but at the same time the radiation balance average is close to  $5 \text{ Wm}^{-2}$ . This suggests that this version of the model leaks energy by the same amount. The other simulation instead exhibits an initial negative radiation balance, which combines with the energy leakage to cause a strong cooling drift. It has proven difficult to locate where in the model this happens, although it is clear that energy is lost in the dynamical core, and also the cloud micro physics contains errors. Work is currently ongoing to remove these bugs.

We encountered an interesting example wherein the radiation imbalance did not cause an immediate temperature drift in the experiment dpp0052 (Figure 4); although the radiation balance is well below that observed by  $5\text{--}10 \text{ Wm}^{-2}$ , the global mean surface temperature is about  $0.5\text{--}1 \text{ K}$  above the observed throughout the simulation. We were able to isolate this behavior to a bug in the momentum forcing which was introduced inadvertently along with an updated vertical coordinate



**Figure 7** Evolution of temperature versus radiation balance in the two longest coupled simulations. Shown is the monthly mean temperature versus radiation balance for the two longer simulations and observations. Shown as thin orange lines are the 20 individual cycles of observations. Dashed orange lines are means of the observations, and as described within the figure we estimate the equilibrium temperature of dpp0029/33 from the drift between the two simulated years.



**Figure 8** Two 10 year long simulations with ICON-Sapphire at 10 km resolution, as well as observations that are also shown in Figure 7. The ngc2012 simulation drifts to colder temperatures with time.

system in the ocean, by reverting only this change in dpp0066 (Table 1). The problem permitted strong stratification in the uppermost meters of the ocean to develop in summer. The result is very large positive surface temperature biases in the summer hemisphere (Figure 5e), explaining the unexpected evolution of the global mean temperature. Eventually, though, the oceans would have to cool in response to the negative radiation balance.

All in all, there are many new challenges associated with tuning the radiation balance and global mean temperature in global coupled cloud resolving models. But they are not insurmountable. Key to this will be to both build up an understanding of the role of the tunable parameters – here we have yet to familiarise ourselves with the new turbulence and cloud micro physics schemes – but also to eliminate structural problems such as energy conservation and other issues such as that encountered here with the ocean vertical mixing. Drifts can, however, not be eliminated entirely and users should be aware of this and take it into account in their analyses. Nevertheless, as the spin ups are short due to computational limitations, so are also the experiments that can be done with the model. Therefore the drifts that can be tolerated are much larger than with contemporary climate models.

#### 2.5.4 Strikingly familiar sea surface temperature biases

Whereas the above has focused mostly on global mean temperature, we would like to also point attention to the distribution of surface temperature biases (Figure 5). First, one can distinguish phase 2 runs with prognostic cloud micro physics which were dominated by warm biased tropical sea surface temperatures (panels a and b), from phase 3 runs with the Smagorinsky turbulence mixing scheme which instead are dominated by cold biased lands (panels c-f). Here dpp0052 (panel e) stands out due to the ocean momentum forcing bug as discussed in the previous section.

Nevertheless, there are also interesting commonalities among the simulations: the warm bias in the tropical eastern boundary up-welling stratocumulus dominated regions, the warm bias in the Southern Ocean and the cold bias in the North Atlantic south of Greenland. All three regions are also commonly biased in contemporary climate models. The warm bias in the stratocumulus regions off the coasts of California, Peru, Namibia and Australia are thought to be a complex problem combining too few clouds with erroneous coastal winds and ocean currents, some of which might be helped by higher resolutions (Zuidema et al., 2016). Apparently, however, they persist at 5 km grid spacing: presumably the coastal jets and ocean currents are well-resolved suggesting that instead the poor representation of stratocumulus clouds which involve finer scales of motion is the main source of error. The Southern Ocean warm bias is mostly

due to issues with clouds being insufficiently reflective (Hyder et al., 2018). The Southern Ocean warm bias is not evident in dpp0001, but since it is analysed in September to November, which is austral spring, it may be tied to sea ice melt in a way that the other runs are not. The other simulations exhibit a clear warm bias, although it should be noted this appears in the southern hemisphere winter months. The North Atlantic cold bias is commonly found in coupled climate models and thought to be related to poorly represented ocean currents (Wang et al., 2014).

It is intriguing that these three long standing climate model biases remain at high resolutions, at least in the case of these early ICON-Sapphire runs, suggesting that either even higher resolutions are necessary or that the remaining physics parameterisations are the culprit. As such the results can help narrow down the causes of biases also in contemporary climate models.

### 3 RESPONSE TO INCREASING CO<sub>2</sub>

It is of particular interest to see how the response to increased atmospheric carbon dioxide (CO<sub>2</sub>) of the less parameterised ICON-Sapphire model compares to contemporary climate models. Although the available run is insufficient to use the long-term response to probe the model's climate sensitivity, it is possible to look at the fast response to CO<sub>2</sub>. The fast precipitation response is a major part of the long term response (Bony et al., 2013), and hence it is interesting whether the modeling approach taken here leads to a different response. Furthermore, sea surface temperature patterns arising from ocean heat uptake are thought to play an important role in setting the transient warming rate (Winton et al., 2010; Held et al., 2010; Armour et al., 2013). Studies have suggested that contemporary coupled climate models underestimate the strength of these patterns, and hence their dampening effect, in particular by warming too fast in the East Pacific (Zhou et al., 2016). If this bias is related to an inability of contemporary climate models to resolve ocean up-welling in the East Pacific, then the response in the ICON-Sapphire runs should exhibit less warming in that region.

To investigate these ideas a four month simulation with ICON-Sapphire with quadrupled CO<sub>2</sub> starting on 1 August (dpp0005) is compared to an ensemble of simulations conducted with MPI-ESM1.2-LR (Mauritsen et al., 2019). This ensemble consists of 10 runs with quadrupled CO<sub>2</sub> also started on 1 August with initial conditions sampled from different years of a pre-industrial control simulation to sample internal variability. The response is then calculated as the difference relative to the control simulation over the corresponding period and/or region. As a result internal variability in both the quadrupled CO<sub>2</sub> experiment and the control impacts the result, compared to common practice where a long control simulation is

averaged to eliminate internal variability. In our case, though, this effect is the same in both the ICON-Sapphire and the MPI-ESM1.2-LR experiments.

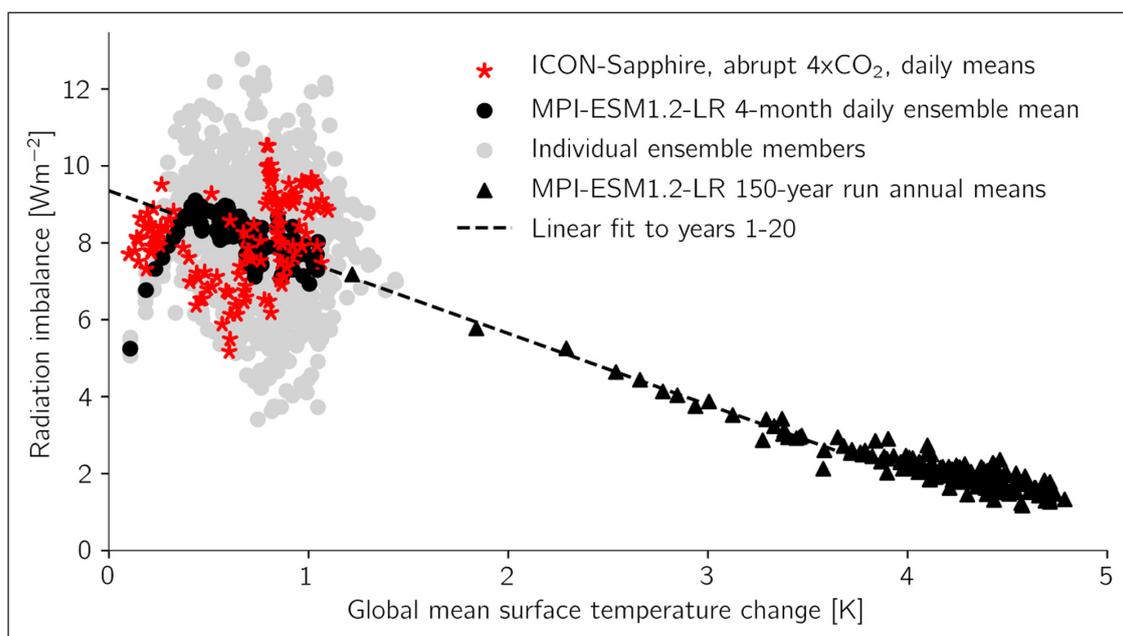
Scattering global means of surface temperature against top-of-atmosphere radiation balance is a common way to estimate forcing and feedback in climate models (Gregory et al., 2004). In the four months simulated here, both ICON-Sapphire and MPI-ESM1.2-LR warm by about 1 K and they exhibit similar radiation balances (Figure 9), as most of the daily means of ICON-Sapphire are within the ten member ensemble. The ensemble mean exhibits a short adjustment with rising radiation imbalance for a couple of weeks, as primarily the stratosphere cools, followed by a slow decline with further warming in line with the expected feedback as estimated by a linear fit to the first 20 years of a longer simulation. Thus, the high resolution ICON-Sapphire simulation global mean fast response is indistinguishable from that of the contemporary MPI-ESM1.2-LR climate model, and potentially the decadal feedback could be studied with just a few years of simulation.

Inspecting next the zonal mean surface temperature and precipitation response reveals a temperature response that is surprisingly similar between the two models (Figure 10). The ICON-Sapphire model is within the ensemble at most latitudes, albeit towards the least warming end in the tropics and sub-tropics south of the Equator and in the northern hemisphere mid-latitudes, and it warms the more than any ensemble member over part of the Antarctic. Zonal mean precipitation change is more variable among the ten MPI-ESM1.2-LR ensemble members, and also here ICON-Sapphire lies mostly within

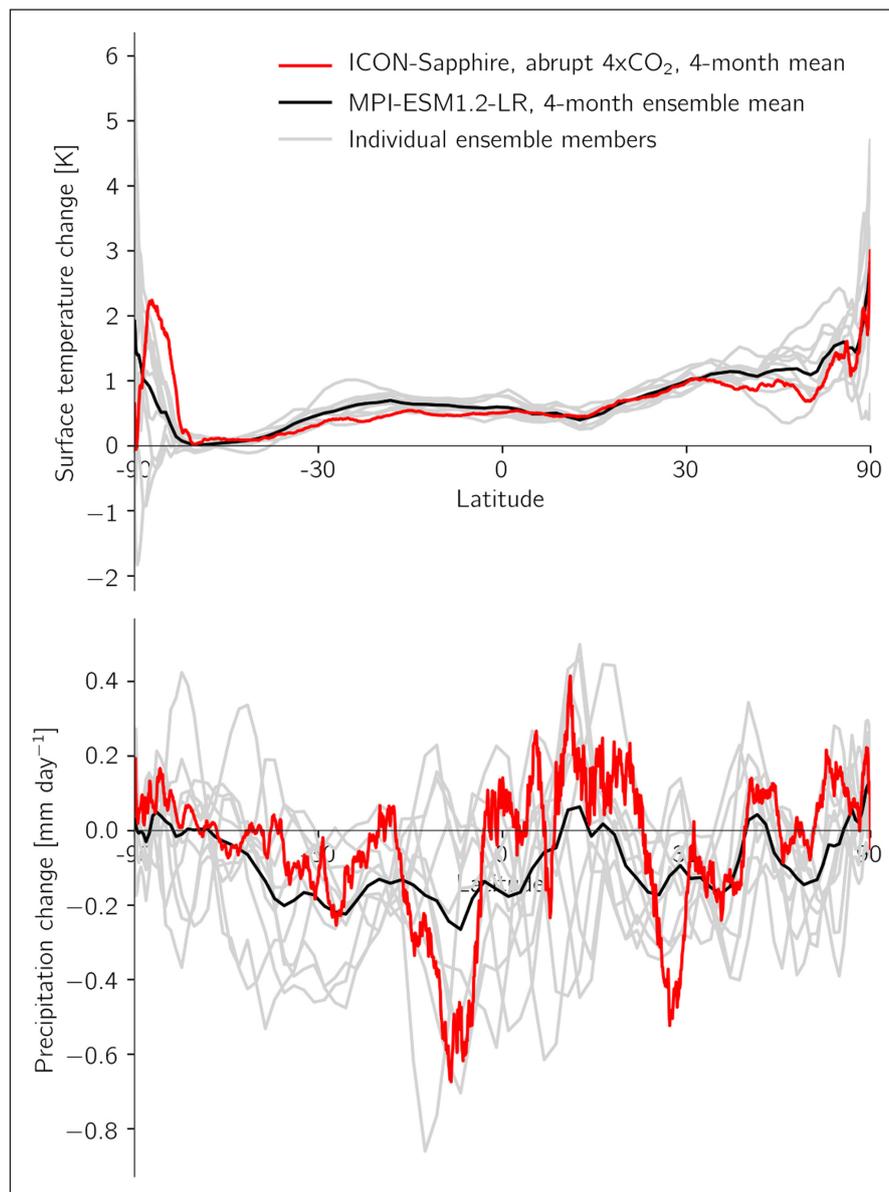
the spread. Nevertheless, there is a systematic pattern with stronger increase of precipitation in the ITCZ region and strong decreases in the sub-tropics in the ICON-Sapphire simulation suggestive of a narrowing of the tropical rain band, even if the model is not an obvious outlier.

Returning to the surface temperature change, also the spatial structure of warming is similar between the MPI-ESM1.2-LR ensemble mean and the ICON-Sapphire response to quadrupled  $\text{CO}_2$  (Figure 11). There are regions where the response is different, for example cooling in the North Atlantic, in terms of the east-west gradient of warming in the tropical Pacific or generally less warming on land, however, it is to be expected that a single realisation contains more noise than an ensemble mean due to internal variability. To investigate this we calculate in each gridpoint the rank of ICON-Sapphire within the MPI-ESM1.2-LR ensemble: if ICON-Sapphire is coldest it is assigned rank 0, if it is the warmest rank 10. This confirms the impression from before, but also the tropical Atlantic stands out with ICON-Sapphire ranking generally the highest. The corresponding rank histogram further shows that the surface temperature pattern of ICON-Sapphire is stronger than that of MPI-ESM1.2-LR with an over representation of ranks 0,1 and 10 (Figure 12). The map of precipitation rank is too noisy to be informative, nevertheless the rank histogram of precipitation shows an over representation foremost of rank 10, which is points with more precipitation in ICON-Sapphire than any of the MPI-ESM1.2-LR ensemble members.

All in all, the general response of radiation, temperature and precipitation to  $\text{CO}_2$  in an early version of ICON-Sapphire is to first order surprisingly similar



**Figure 9** Top of atmosphere radiation balance versus global surface temperature in response to an abrupt quadrupling of atmospheric  $\text{CO}_2$ . Shown as stars and circles are daily means from 4-month simulations starting 1 August in a single realisation with ICON-Sapphire and ten realisations with MPI-ESM1.2-LR, as well as the ensemble mean of the latter. Triangles shows yearly means from a 150-year run. The dashed line is a linear fit to the latter run years 1–20.



**Figure 10** Zonal mean temperature and precipitation change in the first four months following a quadrupling of atmospheric  $\text{CO}_2$ .

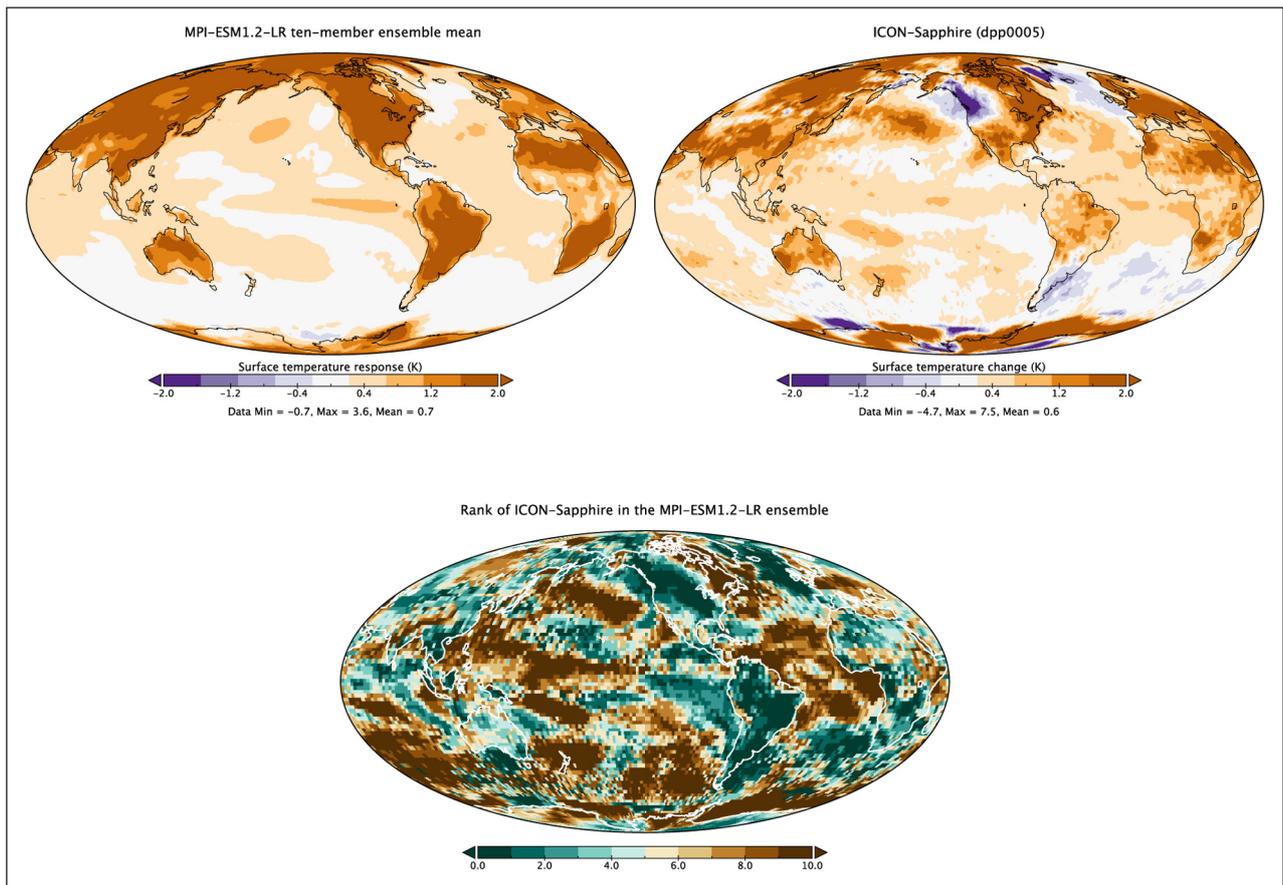
to that in MPI-ESM1.2-LR, given the different nature of the models. There are however intriguing differences such as the stronger surface temperature patterns are interesting in that they could explain some biases common to contemporary climate models (e.g. Zhou et al., 2016), and the results for precipitation suggestive of a sharpening of the ITCZ in a warming world. The use of an ensemble for the lower resolution model appears promising, and suggests that already with a few years simulations with  $\text{CO}_2$  forcing in ICON-Sapphire more firm conclusions regarding feedback mechanisms and other possible differences can be drawn.

#### 4 CONCLUDING REMARKS

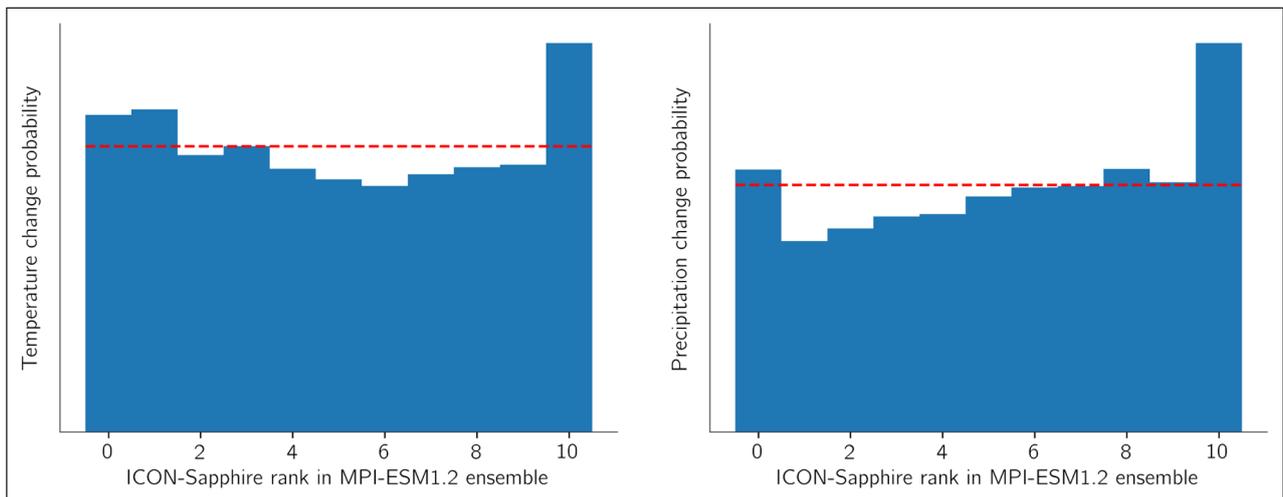
The advances demonstrated here in both developing and utilising a coupled cloud- and ocean eddy-resolving Earth system model represents an important step towards

leveraging exascale computing systems that will emerge in the coming years for weather and climate studies. As we have shown, super computing systems are at the verge of being able to run global simulations at the kilometre scale with a throughput of several months up to nearly a year per day. Thereafter, such simulations will become cheaper, but not much faster. Hence, we argue, the time to develop these models is now.

Unlike previous incremental improvements from climate model resolution increases, the move to kilometre scale resolutions represents a step change. By simulating, rather than parameterising, moist convection, gravity waves and ocean eddies using the equations of motion we are able to make the model codes simpler and the results easier to understand, thereby injecting us with a genuine hope of gaining new insights. Beyond advances in scientific understanding, the implementation of kilometre resolution models at scale can bring climate science much closer to users



**Figure 11** Mean temperature response during first four months following a quadrupling of atmospheric CO<sub>2</sub>. The upper left panel shows the 10-member ensemble mean from the MPI-ESM1.2-LR model and the right panel is from the ICON-Sapphire model, interpolated to the same T63 grid. The lower panel shows the rank of ICON-Sapphire in the MPI-ESM1.2-LR ensemble, whereby zero means it is the coldest and 10 means it is the warmest.



**Figure 12** Rank histogram of temperature (left) and precipitation change (right) in ICON-Sapphire relative to the MPI-ESM1.2-LR ensemble in the first four months following a quadrupling of atmospheric CO<sub>2</sub>, whereby zero means it is the coldest or driest and 10 means it is the warmest or wettest.

of climate change information by acting as so-called digital twins to Earth's weather and climate. For instance the local impacts of climate change, such as extremes of precipitation, storms and droughts which are hardly represented in a meaningful way by current climate models can be simulated directly with such models.

Another way in which these models can be useful is by testing whether they provide results that are out of sample relative to contemporary climate models. We provided such an example with the fast response to quadrupled CO<sub>2</sub>, whereby a single expensive experiment with ICON-Sapphire was compared to a computationally inexpensive

ensemble with the CMIP6 class model MPI-ESM1.2-LR. The purpose of the ensemble is to assess whether the single ICON-Sapphire experiment is within the range of internal variability of MPI-ESM1.2-LR. Our initial investigations, based on an early version of ICON-Sapphire, did not provide strong evidence of out of sample behavior relative to MPI-ESM1.2-LR despite the vast difference in resolution and parameterisation. What we do find is different is that the fast response of surface temperature and precipitation is more diverse in ICON-Sapphire, with a larger representation of both weak and strong warming, as well as increases in strong precipitation.

An important challenge is to limit model drifts to levels that are acceptable for the envisioned purposes of the model, something which is usually done by tuning the radiation balance using parameters related to cloud processes. However, at kilometre scale resolutions the convection parameterisation can be turned off, reducing the number of parameters, and at the same time the parameters which remain are less uncertain. Therefore model developers may face situations where they have to work with parameters outside their estimated range of uncertainty. In our case, fortunately, the radiation balance was already quite close to the observed, and minimal tuning was necessary. Instead, we found in several instances that model physics changes, such as the turbulence mixing schemes in both the ocean and atmosphere, and model energy leakages had large impacts on model drifts. It is our hope that other institutes that are also pursuing global cloud resolving models can benefit from some of our experiences.

## ACKNOWLEDGEMENTS

This work is supported by the Max-Planck-Gesellschaft (MPG). TM acknowledges funding from the European Research Council (ERC) (Grant agreement No.770765) and the European Union's Horizon 2020 research and innovation program projects CONSTRAIN and NextGEMS (Grant agreements No.820829 and No.101003470). Computational resources were made available by Deutsches Klimarechenzentrum (DKRZ) in Hamburg through support from Bundesministerium für Bildung und Forschung (BMBF), by the the Gauss Centre for Supercomputing e.V. on the supercomputer JUWELS-Booster at the Jülich Supercomputing Centre (JSC), and by the Swedish National Infrastructure for Computing (SNIC) at the National Supercomputing Centre (NSC) in Linköping partially funded by the Swedish Research Council through grant agreement no. 2018-05973.

## COMPETING INTERESTS

The authors have no competing interests to declare.

## AUTHOR AFFILIATIONS

**Thorsten Mauritsen**  [orcid.org/0000-0003-1418-4077](https://orcid.org/0000-0003-1418-4077)  
Meteorologiska Institutionen vid Stockholms Universitet (MISU), Stockholm, Sweden

**Rene Redler**  
Max Planck Institut für Meteorologie (MPI-M), Hamburg, Germany

**Monika Esch**  
Max Planck Institut für Meteorologie (MPI-M), Hamburg, Germany

**Bjorn Stevens**  [orcid.org/0000-0003-3795-0475](https://orcid.org/0000-0003-3795-0475)  
Max Planck Institut für Meteorologie (MPI-M), Hamburg, Germany

**Cathy Hohenegger**  [orcid.org/0000-0002-7478-6275](https://orcid.org/0000-0002-7478-6275)  
Max Planck Institut für Meteorologie (MPI-M), Hamburg, Germany

**Daniel Klocke**  [orcid.org/0000-0001-7405-013X](https://orcid.org/0000-0001-7405-013X)  
Max Planck Institut für Meteorologie (MPI-M), Hamburg, Germany

**Renate Brokopf**  
Max Planck Institut für Meteorologie (MPI-M), Hamburg, Germany

**Helmuth Haak**  [orcid.org/0000-0002-9883-5086](https://orcid.org/0000-0002-9883-5086)  
Max Planck Institut für Meteorologie (MPI-M), Hamburg, Germany

**Leonidas Linardakis**  
Max Planck Institut für Meteorologie (MPI-M), Hamburg, Germany

**Niklas Röber**  
Deutsches Klimarechenzentrum GmbH (DKRZ), Hamburg, Germany

**Reiner Schnur**  [orcid.org/0000-0002-7380-8313](https://orcid.org/0000-0002-7380-8313)  
Max Planck Institut für Meteorologie (MPI-M), Hamburg, Germany

## REFERENCES

- Amdahl, GM.** 1967. Validity of single processor approach to achieving large scale computing capabilities. *AFIPS Conference Proceedings*, 30: 483–485.
- Armour, KC, Bitz, CM and Roe, GH.** 2013. Time-Varying Climate Sensitivity from Regional Feedbacks. *Journal of Climate*, 26(13): 4518–4534. DOI: <https://doi.org/10.1175/JCLI-D-12-00544.1>
- Balaji, V,** et al. 2022. Are GCMs obsolete? Manuscript submitted.
- Baldauf, M, Seifert, A, Foerstner, J, Majewski, D, Raschendorfer, M and Reinhardt, T.** 2011. Operational convective-scale numerical weather prediction with the COSMO model: Description and sensitivities. *Monthly Weather Review*, 139(12): 3887–3905. DOI: <https://doi.org/10.1175/MWR-D-10-05013.1>
- Bjerknes, V.** 1904. Das Problem der Wettervorhersage, betrachtet vom Standpunkte der Mechanik und der Physik. *Meteorol. Z.*, 21: 1–7.
- Bolin, B.** 1955. Numerical forecasting with the barotropic model. *Tellus*, 7: 27–49. DOI: <https://doi.org/10.3402/tellusa.v7i1.8770>
- Bony, S, Bellon, G, Klocke, D,** et al. 2013. Robust direct effect of carbon dioxide on tropical circulation and regional precipitation. *Nature Geosci*, 6: 447–451. DOI: <https://doi.org/10.1038/ngeo1799>
- Bony, S, Stevens, B, Ament, F,** et al. 2017. EUREC4A: A Field Campaign to Elucidate the Couplings Between Clouds, Convection and Circulation. *Surv Geophys*, 38: 1529–1568. DOI: <https://doi.org/10.1007/s10712-017-9428-0>

- Charney, JG, Fjörtoft, R and von Neumann, J.** 1950. Numerical integration of the barotropic vorticity equation. *Tellus*, 2(4): 237–254. DOI: <https://doi.org/10.3402/tellusa.v2i4.8607>
- Deardorff, J.** 1970. A numerical study of three-dimensional turbulent channel flow at large Reynolds numbers. *Journal of Fluid Mechanics*, 41(2): 453–480. DOI: <https://doi.org/10.1017/S0022112070000691>
- Dipankar, A, Stevens, B, Heinze, R, Moseley, C, Zängl, G, Giorgetta, M and Brdar, S.** 2015. Large eddy simulation using the general circulation model ICON. *J. Adv. Model. Earth Syst.*, 7: 963–986. DOI: <https://doi.org/10.1002/2015MS000431>
- Fiedler, S, Crueger, T, D'Agostino, R, et al.** 2020. Simulated Tropical Precipitation Assessed across Three Major Phases of the Coupled Model Intercomparison Project (CMIP). *Mon. Wea. Rev.*, 148: 3653–3680. DOI: <https://doi.org/10.1175/MWR-D-19-0404.1>
- Flynn, C and Mauritsen, T.** 2020. On the climate sensitivity and historical warming evolution in recent coupled model ensembles. *Atmos. Chem. Phys.*, 20: 7829–7842. DOI: <https://doi.org/10.5194/acp-20-7829-2020>
- Fuhrer, O, Chadha, T, Hoefler, T, Kwasniewski, G, Lapillonne, X, Leutwyler, D, Lüthi, D, Osuna, C, Schär, C, Schulthess, TC and Vogt, H.** 2018. Near-global climate simulation at 1 km resolution: establishing a performance baseline on 4888 GPUs with COSMO 5.0. *Geosci. Model Dev.*, 11: 1665–1681. DOI: <https://doi.org/10.5194/gmd-11-1665-2018>
- Giorgetta, MA, Brokopf, R, Crueger, T, Esch, M, Fiedler, S, Helmert, J, Hohenegger, C, Kornblueh, L, Köhler, M, Manzini, E, Mauritsen, T, Nam, C, Raddatz, T, Rast, S, Reinert, D, Sakradzija, M, Schmidt, H, Schneck, R, Schnur, R, Silvers, L, Wan, H, Zängl, G and Stevens, B.** 2018. ICON-A, the atmosphere component of the ICON Earth system model: I. Model description. *Journal of Advances in Modeling Earth Systems*, 10: 1613–1637. DOI: <https://doi.org/10.1029/2017MS001242>
- Gregory, JM, Ingram, WJ, Palmer, MA, Jones, GS, Stott, PA, Thorpe, RB, Lowe, JA, Johns, TC and Williams, KD.** 2004. A new method for diagnosing radiative forcing and climate sensitivity. *Geophys. Res. Lett.*, 31: L03205. DOI: <https://doi.org/10.1029/2003GL018747>
- Hanke, M and Redler, R.** 2019. New features with YAC 1.5.0. Reports on ICON, No 3. DOI: [https://doi.org/10.5676/DWD\\_pub/nwv/icon\\_003](https://doi.org/10.5676/DWD_pub/nwv/icon_003)
- Hanke, M, Redler, R, Holfeld, T and Yastremsky, M.** 2016. YAC 1.2.0: new aspects for coupling software in Earth system modelling. *Geosci. Model Dev.*, 9: 2755–2769. DOI: <https://doi.org/10.5194/gmd-9-2755-2016>
- Heikes, RH and Randall, DA.** 1995a. Numerical integration of the shallow- water equations on a twisted icosahedral grid. Part I: Basic design and results of tests. *Mon. Wea. Rev.*, 123: 1862–1880. DOI: [https://doi.org/10.1175/1520-0493\(1995\)123<1862:NIOTSW>2.0.CO;2](https://doi.org/10.1175/1520-0493(1995)123<1862:NIOTSW>2.0.CO;2)
- Heikes, RH and Randall, DA.** 1995b. Numerical integration of the shallow- water equations on a twisted icosahedral grid. Part II: A detailed description of the grid and analysis of numerical accuracy. *Mon. Wea. Rev.*, 123: 1881–1887. DOI: [https://doi.org/10.1175/1520-0493\(1995\)123<1881:NIOTSW>2.0.CO;2](https://doi.org/10.1175/1520-0493(1995)123<1881:NIOTSW>2.0.CO;2)
- Heinze, R, Dipankar, A, Henken, CC, Moseley, C, Sourdeval, O, Trömel, S, Xie, X, Adamidis, P, Ament, F, Baars, H, Barthlott, C, Behrendt, A, Blahak, U, Bley, S, Brdar, S, Brueck, M, Crewell, S, Deneke, H, Di Girolamo, P, Evaristo, R, Fischer, J, Frank, C, Friederichs, P, Göcke, T, Gorges, K, Hande, L, Hanke, M, Hansen, A, Hege, H-C, Hoose, C, Jahns, T, Kalthoff, N, Klocke, D, Kneifel, S, Knippertz, P, Kuhn, A, van Laar, T, Macke, A, Maurer, V, Mayer, B, Meyer, CI, Muppa, SK, Neggers, RAJ, Orlandi, E, Pantillon, F, Pospichal, B, Röber, N, Scheck, L, Seifert, A, Seifert, P, Senf, F, Siligam, P, Simmer, C, Steinke, S, Stevens, B, Wapler, K, Weniger, M, Wulfmeyer, V, Zängl, G, Zhang, D and Quaas, J.** 2017. Large-eddy simulations over Germany using ICON: a comprehensive evaluation. *Q.J.R. Meteorol. Soc.*, 143: 69–100. DOI: <https://doi.org/10.1002/qj.2947>
- Held, IM, Winton, M, Takahashi, K, Delworth, T, Zeng, F and Vallis, GK.** 2010. Probing the Fast and Slow Components of Global Warming by Returning Abruptly to Preindustrial Forcing. *Journal of Climate*, 23(9): 2418–2427. DOI: <https://doi.org/10.1175/2009JCLI3466.1>
- Hohenegger, C, Kornblueh, L, Klocke, D, Becker, T, Cioni, G, Engels, JF, Schulzweida, U and Stevens, B.** 2020. Climate statistics in global simulations of the atmosphere, from 80 to 2.5 km grid spacing. *J. Meteorol. Society Japan*, 98: 73–91. DOI: <https://doi.org/10.2151/jmsj.2020-005>
- Hohenegger, C, Korn, P, Linardakis, L, Redler, R, Schnur, R, Adamidis, P, Bao, J, Bastin, S, Behravesh, M, Bergemann, M, Biercamp, J, Bockelmann, H, Brokopf, R, Bruggemann, N, Esch, M, George, G, Giorgetta, MA, Gutjahr, O, Haak, H, Hanke, M, Jahns, T, Jungclaus, J, Kern, M, Klocke, D, Kluft, L, Kornblueh, L, Kosukhin, SS, Kroll, C, Lee, J, Luschow, V, Mauritsen, T, Müller, R, Naumann, AK, Paccini, L, Panos, S, Praturi, D, Putrasahan, D, Rast, S, Riddick, T, Roeber, N, Schmidt, H, Schulzweida, U, Segura, H, Shevchenko, R, Singh, V, Specht, MS, Stephan, C, von Storch, J-S, Vogel, R, Wengel, C, Winkler, M, Ziemann, F and Stevens, B.** 2022. The ICON Sapphire model: a climate modeling system designed for applications at kilo- and subkilometer scales. In preparation.
- Hohenegger, C and Stevens, B.** 2018. The role of the permanent wilting point in controlling the spatial distribution of precipitation. *Proc. Natl. Acad. Sci.*, 115: 5692–5697. DOI: <https://doi.org/10.1073/pnas.1718842115>
- Hourdin, F, Mauritsen, T, Gettelman, A, Golaz, J, Balaji, V, Duan, Q, Folini, D, Ji, D, Klocke, D, Qian, Y, Rauser, F, Rio, C, Tomassini, L, Watanabe, M and Williamson, D.** 2017. The Art and Science of Climate Model Tuning. *Bulletin of the American Meteorological Society*, 98(3): 589–602. DOI: <https://doi.org/10.1175/BAMS-D-15-00135.1>
- Hyder, P, Edwards, J, Allan, RP, et al.** 2018. Critical Southern Ocean climate model biases traced to atmospheric model cloud errors. *Nature Communication*, 9: 3625. DOI: <https://doi.org/10.1038/s41467-018-05634-2>

- Jungclaus, JH**, et al. 2021. The ICON Earth System Model Version 1.0. *Earth and Space Science Open Archive*. DOI: <https://doi.org/10.1002/essoar.10507989.1>
- Klemp, JB** and **Wilhelmson, RB**. 1978. The Simulation of Three-Dimensional Convective Storm Dynamics. *Journal of Atmospheric Sciences*, 35(6): 1070–1096. DOI: [https://doi.org/10.1175/1520-0469\(1978\)035<1070:TSOTDC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1978)035<1070:TSOTDC>2.0.CO;2)
- Korn, P** and **Linardakis, L**. 2018. A conservative discretization of the shallow-water equations on triangular grid. *J. Comp. Phys.*, 375: 871–900. DOI: <https://doi.org/10.1016/j.jcp.2018.09.002>
- Loeb, NG, Doelling, DR, Wang, H, Su, W, Nguyen, C, Corbett, JG, Liang, L, Mitrescu, C, Rose, FG** and **Kato, S**. 2018. Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Top-of-Atmosphere (TOA) Edition-4.0 Data Product. *J. Climate*, 31: 895–918. DOI: <https://doi.org/10.1175/JCLI-D-17-0208.1>
- Lorenz, EN**. 1969. The predictability of a flow which possesses many scales of motion. *Tellus*, 21(3): 289–307. DOI: <https://doi.org/10.3402/tellusa.v21i3.10086>
- Lott, F**. 1999. Alleviation of stationary biases in a GCM through a mountain drag parameterization scheme and a simple representation of mountain lift forces. *Mon. Wea. Rev.*, 127: 788–801. DOI: [https://doi.org/10.1175/1520-0493\(1999\)127<0788:AOSBIA>2.0.CO;2](https://doi.org/10.1175/1520-0493(1999)127<0788:AOSBIA>2.0.CO;2)
- Manabe, S** and **Bryan, K**. 1969. Climate calculations with a combined ocean-atmosphere model. *Journal of Atmospheric Sciences*, 26(4): 786–789. DOI: [https://doi.org/10.1175/1520-0469\(1969\)026<0786:CCWACO>2.0.CO;2](https://doi.org/10.1175/1520-0469(1969)026<0786:CCWACO>2.0.CO;2)
- Manabe, S, Smagorinski, J** and **Strickler, RF**. 1965. Simulated climatology of a general circulation model with a hydrological cycle. *Monthly Weather Review*, 93(12): 769–798. DOI: [https://doi.org/10.1175/1520-0493\(1965\)093<0769:SCOAGC>2.3.CO;2](https://doi.org/10.1175/1520-0493(1965)093<0769:SCOAGC>2.3.CO;2)
- Mauritsen, T**, et al. 2012. Tuning the climate of a global model. *J. Adv. Model. Earth Syst.*, 4: M00A01. DOI: <https://doi.org/10.1029/2012MS000154>
- Mauritsen, T, Bader, J, Becker, T, Behrens, J, Bittner, M, Brokopf, R**, et al. 2019. Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and its response to increasing CO<sub>2</sub>. *Journal of Advances in Modeling Earth Systems*, 11: 998–1038. DOI: <https://doi.org/10.1029/2018MS001400>
- Moore, GE**. 1965. Cramping more components onto integrated circuits. *Electronics Magazine*, 38.
- Morice, CP, Kennedy, JJ, Rayner, NA, Winn, JP, Hogan, E, Killick, RE, Dunn, RJH, Osborn, TJ, Jones, PD** and **Simpson, IR**. 2021. An updated assessment of near-surface temperature change from 1850: the HadCRUT5 dataset. *Journal of Geophysical Research*. DOI: <https://doi.org/10.1029/2019JD032361>
- Palmer, T** and **Stevens, B**. 2019. The scientific challenge of understanding and estimating climate change. *Proc. Natl. Acad. Sci.*, 116(49): 24390–24395. DOI: <https://doi.org/10.1073/pnas.1906691116>
- Phillips, NA**. 1956. The general circulation of the atmosphere: A numerical experiment. *Q.J.R. Meteorol. Soc.*, 82: 123–164. DOI: <https://doi.org/10.1002/qj.49708235202>
- Pithan, F, Angevine, W** and **Mauritsen, T**. 2015. Improving a global model from the boundary layer: Total turbulent energy and the neutral limit Prandtl number. *J. Adv. Model. Earth Syst.*, 7. DOI: <https://doi.org/10.1002/2014MS000382>
- Radtke, J, Mauritsen, T** and **Hohenegger, C**. 2021. Shallow cumulus cloud feedback in large eddy simulations – bridging the gap to storm-resolving models. *Atmos. Chem. Phys.*, 21: 3275–3288. DOI: <https://doi.org/10.5194/acp-21-3275-2021>
- Retsch, MH, Mauritsen, T** and **Hohenegger, C**. 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: 2070–2088. DOI: <https://doi.org/10.1029/2019MS001677>
- Richardson, LF**. 1922. Weather prediction by numerical process. Cambridge University Press, 258 pp.
- Satoh, M**. 2014. Atmospheric Circulation Dynamics and General Circulation Models. Heidelberg: Springer Verlag. DOI: <https://doi.org/10.1007/978-3-642-13574-3>
- Schneider, T, Lan, S, Stuart, A** and **Teixeira, J**. 2017. Earth system modeling 2.0: A blueprint for models that learn from observations and targeted high-resolution simulations. *Geophysical Research Letters*, 44: 12396–12417. DOI: <https://doi.org/10.1002/2017GL076101>
- Shepherd, T**. 2014. Atmospheric circulation as a source of uncertainty in climate change projections. *Nature Geosci*, 7: 703–708. DOI: <https://doi.org/10.1038/ngeo2253>
- Smith, RD, Maltrud, ME, Bryan, F** and **Hecht, MW**. 2000. Numerical simulation of the North Atlantic Ocean at 1/10°. *Journal of Physical Oceanography*, 30: 1532–1561. DOI: [https://doi.org/10.1175/1520-0485\(2000\)030<1532:NSOTNA>2.0.CO;2](https://doi.org/10.1175/1520-0485(2000)030<1532:NSOTNA>2.0.CO;2)
- Stevens, B**, et al. 2021. EUREC4A. *Earth Syst. Sci. Data*, 13: 4067–4119. DOI: <https://doi.org/10.5194/essd-13-4067-2021>
- Stevens, B, Satoh, M, Auger, L**, et al. 2019. DYAMOND: the Dynamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains. *Prog Earth Planet Sci.*, 6: 61. DOI: <https://doi.org/10.1186/s40645-019-0304-z>
- Sundqvist, H, Berge, E** and **Kristjánsson, JE**. 1989. Condensation and Cloud Parameterization Studies with a Mesoscale Numerical Weather Prediction Model. *Monthly Weather Review*, 117(8): 1641–1657. DOI: [https://doi.org/10.1175/1520-0493\(1989\)117<1641:CACPSW>2.0.CO;2](https://doi.org/10.1175/1520-0493(1989)117<1641:CACPSW>2.0.CO;2)
- Tiedtke, M**. 1989. A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Mon. Wea. Rev.*, 117: 1779–1800. DOI: [https://doi.org/10.1175/1520-0493\(1989\)117<1779:ACMFSF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2)
- Tomita, H, Miura, H, Iga, S, Nasuno, T** and **Satoh, M**. 2005. A global cloud-resolving simulation: Preliminary results from an aqua planet experiment. *Geophys. Res. Lett.*, 32(8): 3283. DOI: <https://doi.org/10.1029/2005GL022459>

- Tomita, H, Satoh, M and Goto, K.** 2002. An optimization of icosahedral grid modified by spring dynamics. *J. Comp. Phys.*, 183: 307–331. DOI: <https://doi.org/10.1006/jcph.2002.7193>
- Tomita, H, Tsugawa, M, Satoh, M and Goto, K.** 2001. Shallow water model on a modified icosahedral geodesic grid by using spring dynamics. *J. Comp. Phys.*, 174: 579–613. DOI: <https://doi.org/10.1006/jcph.2001.6897>
- Uribe, A, Vial, J and Mauritsen, T.** 2021. Sensitivity of tropical extreme precipitation to surface warming in aquaplanet experiments using a global nonhydrostatic model. *Geophysical Research Letters*, 48: e2020GL091371. DOI: <https://doi.org/10.1029/2020GL091371>
- Wan, H, Rasch, PJ, Zhang, K, Qian, Y, Yan, H and Zhao, C.** 2014. Short ensembles: an efficient method for discerning climate-relevant sensitivities in atmospheric general circulation models. *Geosci. Model Dev.*, 7: 1961–1977. DOI: <https://doi.org/10.5194/gmd-7-1961-2014>
- Wang, C, Zhang, L, Lee, SK, Wu, L and Mechoso, CR.** 2014. A global perspective on CMIP5 climate model biases. *Nature Climate Change*, 4: 201–205. DOI: <https://doi.org/10.1038/nclimate2118>
- Washington, WM, Buja, L and Craig, A.** 2008. The computational future for climate and Earth system models: on the path to petaflop and beyond. *Phil. Trans. R. Soc. A.*, 367833–846. DOI: <https://doi.org/10.1098/rsta.2008.0219>
- Weller, H, Weller, HG and Fournier, A.** 2009. Voronoi, Delaunay, and Block-Structured Mesh Refinement for Solution of the Shallow-Water Equations on the Sphere. *Mon. Wea. Rev.*, 137: 4208–4224. DOI: <https://doi.org/10.1175/2009MWR2917.1>
- Williams, KD, Bodas-Salcedo, A, Déqué, M, Fermepin, S, Medeiros, B, Watanabe, M, Jakob, C, Klein, SA, Senior, CA and Williamson, DL.** 2013. The Transpose-AMIP II Experiment and Its Application to the Understanding of Southern Ocean Cloud Biases in Climate Models. *Journal of Climate*, 26(10): 3258–3274. DOI: <https://doi.org/10.1175/JCLI-D-12-00429.1>
- Winton, M, Takahashi, K and Held, IM.** 2010. Importance of Ocean Heat Uptake Efficacy to Transient Climate Change. *Journal of Climate*, 23(9): 2333–2344. DOI: <https://doi.org/10.1175/2009JCLI3139.1>
- Yashiro, H, Terai, M, Yoshida, R, Iga, S-I, Minami, K and Tomita, H.** 2016. Performance Analysis and Optimization of Nonhydrostatic ICosahedral Atmospheric Model (NICAM) on the K Computer and TSUBAME2.5. In: *Proceedings of the Platform for Advanced Scientific Computing Conference on ZZZ – PASC'16*. ACM Press. DOI: <https://doi.org/10.1145/2929908.2929911>
- Zelinka, MD, Myers, TA, McCoy, DT, Po-Chedley, S, Caldwell, PM, Ceppi, P, et al.** 2020. Causes of higher climate sensitivity in CMIP6 models. *Geophysical Research Letters*, 47: e2019GL085782. DOI: <https://doi.org/10.1029/2019GL085782>
- Zhou, C, Zelinka, M and Klein, S.** 2016. Impact of decadal cloud variations on the Earth's energy budget. *Nature Geosci.*, 9: 871–874. DOI: <https://doi.org/10.1038/ngeo2828>
- Zuidema, P, Chang, P, Medeiros, B, Kirtman, BP, Mechoso, R, Schneider, EK, Toniazzo, T, Richter, I, Small, RJ, Bellomo, K, Brandt, P, de Szoeki, S, Farrar, JT, Jung, E, Kato, S, Li, M, Patricola, C, Wang, Z, Wood, R and Xu, Z.** 2016. Challenges and Prospects for Reducing Coupled Climate Model SST Biases in the Eastern Tropical Atlantic and Pacific Oceans: The U.S. CLIVAR Eastern Tropical Oceans Synthesis Working Group. *Bulletin of the American Meteorological Society*, 97(12): 2305–2328. DOI: <https://doi.org/10.1175/BAMS-D-15-00274.1>

---

#### TO CITE THIS ARTICLE:

Mauritsen, T, Redler, R, Esch, M, Stevens, B, Hohenegger, C, Klocke, D, Brokopf, R, Haak, H, Linardakis, L, Röber, N and Schnur, R. 2022. Early Development and Tuning of a Global Coupled Cloud Resolving Model, and its Fast Response to Increasing CO<sub>2</sub>. *Tellus A: Dynamic Meteorology and Oceanography*, 74(2022): 346–363. DOI: <https://doi.org/10.16993/tellusa.54>

**Submitted:** 22 April 2022    **Accepted:** 16 August 2022    **Published:** 02 September 2022

#### COPYRIGHT:

© 2022 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

*Tellus A: Dynamic Meteorology and Oceanography* is a peer-reviewed open access journal published by Stockholm University Press.

