EARLY ONLINE RELEASE

This is a PDF of a manuscript that has been peer-reviewed and accepted for publication. As the article has not yet been formatted, copy edited or proofread, the final published version may be different from the early online release.

This pre-publication manuscript may be downloaded, distributed and used under the provisions of the Creative Commons Attribution 4.0 International (CC BY 4.0) license. It may be cited using the DOI below.

The DOI for this manuscript is
DOI:10.2151/jmsj.2020-005

J-STAGE Advance published date: November 10th, 2019

The final manuscript after publication will replace the preliminary version at the above DOI once it is available.
Climate statistics in global simulations of the atmosphere, from 80 to 2.5 km grid spacing

Cathy Hohenegger

Max Planck Institute for Meteorology, Hamburg, Germany

Luis Kornblueh

Max Planck Institute for Meteorology, Hamburg, Germany

Daniel Klocke

Hans Ertel Centre for Weather Research, Deutscher Wetterdienst, Offenbach, Germany

Tobias Becker

Max Planck Institute for Meteorology, Hamburg, Germany

Guido Cioni

Max Planck Institute for Meteorology, Hamburg, Germany

Jan Frederik Engels

Deutsches Klimarechenzentrum, Hamburg, Germany
Uwe Schulzweida

Max Planck Institute for Meteorology, Hamburg, Germany

Bjorn Stevens

Max Planck Institute for Meteorology, Hamburg, Germany

October 22, 2019

Corresponding author: Cathy Hohenegger, Max Planck Institute for Meteorology, Hamburg, Germany, Bundesstrasse 53 20146 Hamburg.
E-mail: cathy.hohenegger@mpimet.mpg.de
Abstract

Basic climate statistics, such as water and energy budgets, location and width of the InterTropical Convergence Zone (ITCZ), trimodal tropical cloud distribution, position of the polar jet and land-sea contrast remain either biased in coarse-resolution General Circulation Models or are tuned. Here we examine the horizontal resolution dependency of such statistics in a set of global convection-permitting simulations integrated with the ICOsahedral Non-hydrostatic (ICON) model, explicit convection and grid spacings ranging from 80 km down to 2.5 km. The impact of resolution is quantified by comparing the resolution-induced differences to the spread obtained in an ensemble of eight distinct global storm-resolving models.

Using this metric, we find that, at least by 5 km, the resolution-induced differences become smaller than the spread in 26 out of the 27 investigated statistics. Even for 9 (18) of these statistics, a grid spacing of 80 (10) km does not lead to significant differences. Resolution down to 5 km matters especially for net shortwave radiation, which systematically increases with resolution due to reductions in low cloud amount over the subtropical oceans. Further resolution dependencies can be found in the land-to-ocean precipitation ratio, in the latitudinal position and width of the Pacific ITCZ and in the longitudinal position of the Atlantic ITCZ. Also in the tropics, the deep convective cloud population systematically increases at the ex-
pense of the shallow one, whereas the partition of congestus clouds remains fairly constant. Finally, refining the grid spacing systematically moves the simulations closer to observations, but climate statistics exhibiting weaker resolution dependencies are not necessarily associated with smaller biases.
Keywords Climate statistics, General Circulation Models, tropical climate, convection-permitting simulations, convective parameterization

1. Introduction

General Circulation Models (GCMs) are complex tools embodying physical principles to represent the statistics of the climate system. Their application to long time scales and the whole Earth makes the use of resolution fine enough to explicitly represent the major modes of heat transfer challenging. Biases resulting from the use of parameterizations, in particular convective parameterizations, have been persistent (Flato et al. 2013). Several of the basic properties of the climate system that a GCM ought to be able to represent, like width and location of the Intertropical Convergence Zone (ITCZ), water and energy budgets, surface temperature, land-sea contrast, even position of the storm tracks, often exhibit large biases or, if not, it is difficult to assess the extent to which the goodness of fit is a result of error compensation by tuning (Hourdin et al. 2017). Refining the horizontal grid spacing down to a few kilometers and explicitly representing convection, as in so-called convection-permitting models, can bypass at least the problem of having to parameterize deep convection. It nevertheless remains unclear which horizontal resolution is needed to capture basic properties of the climate system.
The use of convection-permitting models is now well established for limited-area weather forecasting (Mass et al. 2002; Richard et al. 2007) and is gaining popularity for regional climate modelling (Prein et al. 2015). Turning off the deep convection scheme, although not the solution to all problems, leads to numerous improvements. The most notable ones are a more realistic timing of the precipitation diurnal cycle (Hohenegger et al. 2008), a higher variability at smaller wavelengths and a better representation of extreme precipitation (Prein et al. 2015; Ban et al. 2014; Chan et al. 2013), as well as a more realistic distribution of precipitation objects (Prein et al. 2013; Wernli et al. 2008; Roberts and Lean 2008). Convection-permitting models allow for the organization and propagation of convective storms (Marsham et al. 2013; Weisman et al. 2008), a notoriously difficult task for convective parameterizations. They also benefit from the finer resolution of their external fields, in particular orography (Prein et al. 2016), which affects the representation of precipitation and of snowpack (Rasmussen et al. 2011; Ikeda et al. 2010). Finally, besides exhibiting an altered precipitation distribution, regional convection-permitting simulations have been found to produce less cloud cover in convective situations due to changes in their cloud morphology, leading to larger net shortwave radiation at the surface as compared to their parameterized counterpart (Prein et al. 2015).
On such regional scales, the convergence behavior depends on the statistics considered. Based on the results of idealized squall line experiments (Weisman et al. 1997), a grid spacing of 4 km was found necessary to represent non-hydrostatic dynamics and to avoid grid-point storms. Whereas properties of individual convective updrafts, like area, velocity, require finer grid spacing than 4 km (Petch et al. 2002; Bryan et al. 2003; Craig and Dörnbrack 2008; Jeevanjee 2017; Hanley et al. 2014), bulk properties, such as domain-averaged precipitation amount or net heating and moistening, have been found to already converge around 4 km (Langhans et al. 2012; Schwartz et al. 2009; Panosetti et al. 2018). Moreover, properties of individual convective updrafts, even if not converged, do not always project strongly on the skill of a simulation (Ito et al. 2017).

Global statistics cannot be derived from such regional studies. Until recently, only the Non-hydrostatic ICosahedral Atmospheric Model (NICAM; Tomita et al. (2005), Miura et al. (2007), Satoh et al. (2008)) had been run at convection-permitting resolution on a global scale, with grid spacings ranging from 14 km down to 0.87 km. Finer grid spacing leads to a later triggering of convection and to a later peak of precipitation, in better agreement with observations (Sato et al. 2008, 2009; Noda et al. 2012). Yashiro et al. (2016) concluded that a minimum grid spacing of 2 to 3 km is necessary to capture the main characteristics of the diurnal cycle. Likewise, continu-
gent upon the chosen definition of deep convective cores, the properties of the simulated convective cores were found to qualitatively change between a grid spacing of 3.5 and 1.7 km (Miyamoto et al. 2013, 2015; Kajikawa et al. 2016; Yashiro et al. 2016). In terms of global statistics, Kajikawa et al. (2016) found no significant resolution dependency for precipitation, vertical mass flux and zonal wind. In contrast, the fractional coverage of different cloud types exhibited significant resolution dependencies, and outgoing longwave radiation became significantly larger when refining the grid spacing from 3.5 km to 1.7 km.

The goal of this study is to examine the horizontal resolution dependencies of basic statistics of the climate system and thus to assess which horizontal grid spacing is required to capture basic statistics of the climate system. We restrict ourselves to a set of statistics that a GCM ought to be able to correctly represent and which are often tuned towards observations in state-of-the-art GCMs. Those concern water and energy budgets, location and width of the ITCZ, cloud distribution in the tropics, as well as jet positions in the extratropics. Moreover, the ability to distinguish between land and ocean will be considered. To assess the resolution dependencies, the latter are compared to the spread of the Dynamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains (DYAMOND) ensemble (Stevens et al. 2019). DYAMOND is an intercom-
parison project of global storm-resolving models, currently comprising nine
models and integrated at a grid spacing of $O(3 \text{ km})$. The motivation for
using the DYAMOND ensemble as a reference dataset is twofold. First,
it allows an objective quantification of the resolution dependency, whereas
traditional convergence studies have to, often subjectively, decide when dif-
ferences between resolutions become small enough. Second, even if variables
have converged, this does not necessarily imply small biases, a point that
will be investigated in this study as well.

Our method involves successively halving the horizontal grid spacing of
global simulations conducted with the ICOsahedral Non-hydrostatic model
(INO; Zängl et al. (2015)), from 80 km down to 2.5 km, following the
DYAMOND experimental protocol. Changes in grid spacing imply changes
in time step and resolution of the external parameter fields, but parameteri-
izations and other parameter settings are kept untouched to be able to assess
the direct effect of resolution changes. This means that all our simulations
use explicit convection, even when run at a grid spacing of 80 km. Even
if this appears as counter-intuitive, past studies (Webb et al. 2015; Maher
et al. 2018) have shown that coarse-resolution models can run stably with-
out using a convection scheme and produce a precipitation climatology that
captures the observed large-scale features. We also only consider the effect
of horizontal resolution, not of vertical resolution. Our study expands on
previous studies based on NICAM, in particular Kajikawa et al. (2016), by considering a larger range of resolutions, partly distinct statistics and using the newly available DYAMOND dataset as a reference to objectively quantify resolution dependencies.

2. Method

2.1 Model

We use the ICON model in a configuration similar to the one employed by the German Weather Service (DWD) for their operational global weather forecasts. Our configuration differs from the latter configuration in not making use of the convection and gravity wave drag parameterizations, in including an additional prognostic variable in the microphysics scheme (graupel) and in calculating the radiation at every grid point. These changes are motivated by the targeted finest resolution of our simulations, namely 2.5 km, whereas the operational weather forecasts are currently run using a grid spacing of 13 km.

In more detailed terms, we employ the ICON model version 2.1.02. Given the targeted resolution of 2.5 km and the DYAMOND experiment protocol (see Section 2.2), the model version had to be slightly updated. Those updates concerned changes for the initialization of humidity, changes
for output data compression and changes for sea surface temperature and sea ice concentration reading. Physical parameterizations in ICON 2.1.02 include the representation of turbulent mixing with a turbulent kinetic energy scheme; a bulk microphysics scheme that predicts cloud water, rain water, cloud ice, snow and graupel (Baldauf et al. 2011); as well as an interactive surface flux scheme and soil model (Schrodin and Heise 2002). Radiative transfer is calculated at every grid point every 15 min using the RRTM (Rapid Radiative Transfer Model) scheme (Mlawer et al. 1997; Mlawer and Clough 1998). Diagnostic fractional cloud cover is calculated with a simple box probability distribution function at every radiation time step. Although the ICON model is currently not used operationally at a grid spacing of 2.5 km, the chosen physical parameterizations stem from the COnsortium for Small-scale MOdeling (COSMO) model, which has been widely used at such fine resolutions, from limited-area operational weather forecasts in the mid-latitudes up to near-global climate simulations (Fuhrer et al. 2018). A similarly configured ICON model version has also been recently used for limited-area simulations over the tropical Atlantic in Klocke et al. (2017).

2.2 Experimental set-up

To investigate horizontal resolution dependencies, we analyse a set of simulations where the horizontal grid spacing is successively refined from
80 km down to 2.5 km by factors of two (see Table 1). The grid spacing corresponds to the square root of the mean cell area of the model triangles. In ICON terminology (see e.g. Giorgetta et al. 2018), the considered grid spacings are R2B5, R2B6, R2B7, R2B8, R2B9 and R2B10. Strictly speaking, all our simulations are convection-permitting as none of them employ a convective parameterization, neither for shallow nor for deep convection.

Ninety levels are used in the vertical with the model top at 75 km. Damping starts in the seventy-seventh layer, above 44 km. The experimental configuration follows the protocol of the DYAMOND model intercomparison project (Stevens et al. 2019). Simulations of 40 days are initialised from the analysis of the atmospheric state from the European Centre for Medium Range Weather Forecasts (ECMWF) on the 1st of August 2016 at 00UTC. Daily observed sea surface temperatures and sea ice cover are prescribed at the bottom boundary. The analysis is available at a grid spacing of 9.5 km.

As we are interested in the direct effect of resolution changes, the physical parameterizations and the model parameters are not tuned to a specific grid spacing. Even the 2.5-km simulation was not tuned for that particular grid spacing. It simply employs the same parameter settings as the operational 13-km ICON model despite using a slightly different set of physical parameterizations (see section 2.1). The only two aspects of the simulations that are adapted as a function of the grid spacing are the model time step
(see Table 1) and the bottom boundary conditions. For the finest resolution, we achieve a stable simulation with a model time step of 22.5 s, a time step that could be successively doubled for the simulations with a grid spacing of 5 and 10 km. For the remaining coarser-resolution simulations, the time step had to be increased by less than double to keep the simulations stable. For each simulated resolution, data for the bottom boundary conditions is recreated by aggregating observations to the used model grid. The resolution of these observational datasets is finer than 2.5 km except for soil texture (0.083°), normalized differential vegetation index (0.4167°), climatological mean near-surface temperature (0.5°), aerosol optical properties (1°), soil albedos for dry or saturated soils (0.5°) and remaining albedo values (0.083°), see Tables 1 and 2 in Asensio et al. (2018).

2.3 Computational aspects

The simulations have been integrated on the supercomputer Mistral of the Deutsches Klimarechenzentrum (DKRZ). The compute partitions comprise 3300 dual socket Intel-CPU nodes with around 265 Terabytes of main memory. Some numbers on model performance are listed in Table 1. They reveal that for the chosen number of nodes, ICON scales well with increases in wall clock time between subsequent grid spacing refinements being less than what would be expected from increases in number of grid points and
time step weighted by the number of nodes. This reflects work that has been conducted in the High Definition Clouds and Precipitation for advancing Climate Prediction (HD(CP)$^2$) project, whose goal was to improve our understanding of cloud and precipitation processes by conducting ICON simulations at $O(100 \text{ m})$ grid spacing over Germany. This effort ensured that ICON scales well (Heinze et al. 2016). The most challenging simulation with a grid spacing of 2.5 km could perform 6 simulated days per day on 540 nodes and is also comparable to what was achieved in DYAMOND by other models (Stevens et al. 2019).

Major problems for the finest-resolution simulation were not per se the computing time, but the generation of input data. As we come close to the resolution of the originating datasets, our simple aggregation procedure did not work anymore and the interpolation procedure had to be updated. The second major problem was the size of the generated output which required updates in the software analysis tool, the Climate Data Operator (CDO) developed at the Max Planck Institute for Meteorology, see discussion in Stevens et al. (2019).

### 2.4 Analysis methodology

To objectively quantify resolution dependencies, we will compare the resolution differences obtained in our ICON simulations to the spread de-
rived from the DYAMOND ensemble of storm-resolving models. Note that, in contrast to past literature, we use here the term “storm-resolving” and not “convection-permitting” to characterize atmospheric models employing grid spacings of a few kilometers (see Satoh et al. (2019) for a discussion of nomenclature). This is so because such models distinguish themselves by their ability to resolve convective storms. Moreover, convection-permitting is an ambiguous term in the context of our study as even our ICON simulation with a grid spacing of 80 km permits convection. It doesn’t use a convective parameterization but still produces convection in the tropics. Except indicated otherwise, the following models of the DYAMOND ensemble are included to compute the spread (see Stevens et al. 2019): ARPEGE-NH (grid spacing of 2.5 km), FV3 (3.3 km), GEOS (3.3 km), ICON (2.5 km), IFS (4.8 km), MPAS (3.8 km), NICAM (3.5 km) and SAM (4.3 km). We do not include the UM model, another model participating in DYAMOND, given its coarser grid spacing of 7.8 km at the equator, which already places it between our ICON simulations with 5 and 10 km grid spacings.

The spread in the DYAMOND ensemble is computed as the standard deviation. It is interpreted as being mainly a result of different physical parameterizations and dynamical cores, whereas the spread in our ICON ensemble of simulations is a result of changes in grid spacing. As such, as long as the difference between a particular grid spacing and the finest
grid spacing of 2.5 km remains smaller than the spread of the DYAMOND ensemble, the resolution is causing differences that are smaller than those associated with the remaining parameterizations and dynamical core. Any of these grid spacings may be seen as appropriate to capture the investigated statistics and one might not need to invest in the high computational burden associated with running simulations at 2.5-km grid spacing. This approach will allow us to determine which grid spacing, between 80 and 2.5 km, is needed to capture basic statistics of the climate system.

This last statement is conditioned on the way we assess resolution differences. As in studies that have looked at convergence, we compare our simulations to the finest available grid spacing which, for computational reason, is 2.5 km. Further refining the grid spacing could increase the calculated resolution differences, in particular for climate statistics that exhibit strong resolution dependencies and that haven’t converged yet. This point is investigated in more details in section 4. Still, the advantage of comparing the resolution-induced differences to the DYAMOND ensemble spread is that it provides us with a clear criterion to decide as to the importance of resolution. Note also that most of the DYAMOND models were run for the first time in such a configuration. Even though the simulations appear to be able to reproduce basic aspects of the observations very well (Stevens et al. 2019) and even though none of them stand out as a clear outlier, they
are error-prone. On the one hand, this means that the spread may be larger than expected. On the other hand, the ICON simulations were not retuned for a specific grid spacing, so that the obtained resolution dependencies may also be larger than expected. In this sense, the two effects can partly cancel out.

We also assess resolution dependencies by looking at 40-day mean statistics of the climate system, as imposed by the experimental protocol of DYAMOND. This bears the risk that resolution differences may not have settled yet. Years long of past experience with tuning the global statistics of a low-resolution GCM (Mauritsen et al. 2012) have nevertheless revealed that short-term integrations, as short as one month, are actually sufficient to assess biases in basic climate properties, such as water and energy budgets. This is in agreement with the results of other studies (Phillips et al. 2004; Sexton et al. 2019) that have found a good match between errors on weather and climatic timescales. This assumption is further confirmed here by looking at the temporal evolution of the resolution-induced differences of the water and energy budgets over the simulated 40 days. No drift could be detected. Taking the net shortwave radiation as an example, which is the variable that will show the largest resolution dependency (see Section 3.1), we obtain global mean differences of -38, -37, -37 and -41 W m\(^{-2}\) between the 80-km and the 2.5-km simulation as averaged over consecutive 10-day
3. Result

3.1 Water and energy budgets

Figure 1 shows mean values of the components of the water and energy budget at the surface and at the top of the atmosphere as a function of the grid spacing of the ICON simulations and expressed as difference to the finest grid spacing. The values are compared to the corresponding standard deviations derived from the DYAMOND ensemble and indicated by the vertical bars in Figure 1.

Given our metric to assess resolution differences, Figure 1 reveals that a grid spacing of 5 km is sufficient to capture the global mean statistics of the water and energy budgets. For all the components of the water and energy budgets, at least by 5 km, the resolution-induced differences become smaller than the DYAMOND spread. For precipitation, sensible heat flux and outgoing longwave radiation, even with a grid spacing of 80 km, the resolution-induced differences remain smaller than the DYAMOND spread. In contrast, resolution strongly affects net shortwave radiation, both at the surface and at the top of the atmosphere, and to a lesser extent net longwave radiation at the surface. Refining the grid spacing from 80 km down to 2.5
km leads to a systematic increase in net shortwave radiation of about 40 W m$^{-2}$ (Figure 1b,c), whereas surface net longwave radiation, with a maximum difference of 10 W m$^{-2}$, gets systematically more negative (Figure 1b). Splitting the surface net longwave radiation in its two components reveals values comprised between -407.3 W m$^{-2}$ and -407.9 W m$^{-2}$ for the emitted longwave radiation across resolutions versus values comprised between 350.7 W m$^{-2}$ and 360.9 W m$^{-2}$ for the downward component. This means that the differences in surface net longwave radiation come solely from its downward component. The latter amounts to 360.9, 358.9, 356.7, 354., 352.3 and 350.7 W m$^{-2}$ at 80, 40, 20, 10, 5 and 2.5 km and hence decreases with finer grid spacings. The similarity of the longwave radiation emitted by the surface across resolutions implies very similar surface temperatures. It mostly reflects the fact that the surface temperature is fixed for 70% of the area by prescribing the sea surface temperature (SST).

In terms of spatial distribution, the largest differences in surface net shortwave radiation can be found over the southern tropical Atlantic, the southeastern tropical Pacific, East of Madagascar reaching almost over to Australia and over the northeastern subtropical Pacific, as shown in Figure 2a using the difference between the coarsest and finest resolution as a representative example. These various regions are all regions that also exhibit large differences in cloud cover (Figure 2b). As those regions are prone to
shallow cumulus or stratocumulus clouds (Medeiros and Stevens 2011), we conclude that changes in resolution mostly affect the representation of such clouds. The 80-km simulation produces much larger cloud amount than the 2.5-km simulation (see Figure 2b). The tendency for the coarser-resolution simulation to produce more cloud cover over subtropical oceanic region than the finer-resolution simulation remains true when considering the range of investigated resolutions. The cloud cover averaged from 20°S to 0°S over oceanic areas only, which encompasses the previously mentioned areas of strong cloud cover difference, amounts to 84%, 76%, 68%, 62%, 58% and 53% at a grid spacing of 80, 40, 20, 10, 5 and 2.5 km, respectively. The cloud cover systematically decreases with finer grid spacing, which is consistent with and explains the systematic increases in net shortwave radiation at the surface, at the top of the atmosphere, and the decrease in downward longwave radiation at the surface, as displayed by Figure 1. The cloud liquid water content exhibits a similar behavior over those regions. The decrease in the surface downward longwave radiation with finer grid spacing could also be partly caused by a cooling and drying of the subcloud layer. Inspection of the temperature field nevertheless reveals the opposite tendency (see e.g., Fig 3a), whereas changes in specific humidity are mixed with some areas exhibiting a drying and other ones a moistening.

The large difference in cloud cover across the simulations already settles
in after 12 hours of simulation. To better understand these differences, we look at profiles of temperature, specific humidity and cloud water for the two extreme simulations under such conditions (Figure 3). The profiles reveal the well-observed structure of the marine boundary layer with the sub-cloud (or mixed) layer below 500-700 m, the cloud layer populated by shallow cumuli (500-2000 m) and the trade inversion around 2 km. The 80-km simulation is associated with a deeper mixed layer (see Figure 3b) as compared to the 2.5-km simulation. This translates itself in drier conditions in the mixed layer but moister and colder conditions in the cloud layer (Figure 3a,c). As a consequence, the cloud layer is more likely to saturate leading to the widespread cloud formation at 80 km (Figure 3c). The deeper mixed layer at 80 km is consistent with a weaker mass transport by shallow cumulus (Neggers et al. 2007), as expected from its poorer representation. Noda et al. (2010) also observed an increase in cloud cover in their NICAM simulations going from 7 to 14 km. They concluded that the subgrid-scale cloud parametrization strongly controlled the low-cloud amount and led to excessive cloud amount at 14 km. In our simulations, the vast majority of the clouds are the result of the full saturation of grid cells rather than of subgrid-scale nature.

The representation of shallow cumulus convection and stratocumulus is expected to have a very small impact on outgoing longwave radiation as the
temperature contrast between the surface is small. This is consistent with
the negligible sensitivity of outgoing longwave radiation to resolution as
displayed by Figure 1c. Moreover and although the simulations with higher
horizontal resolutions have less cloud cover in the subtropics and hence
larger areas with clear sky, meaning higher outgoing longwave radiation,
this is compensated by more extended and colder anvils in the region with
deep convection. Within the main ITCZ region (5°-15° N), the outgoing
longwave radiation is getting less negative with finer grid spacings. The
values are -247, -248, -246, -243, -241, -238 W m \(^{-2}\) at 80, 40, 20, 10, 5 and
2.5 km.

The more modest resolution dependencies of precipitation, sensible heat
flux and latent heat flux, as compared to the radiative fluxes (Figure 1a,b),
can be mostly understood from energy constraints. The changes in net
shortwave radiation at the surface and at the top of the atmosphere almost
compensate (compare Figures 1b and c). Given the decrease in downward
longwave radiation at the surface at finer resolutions, the atmosphere ra-
diative cooling, that needs to be balanced by the surface fluxes, decreases.
The radiative cooling, for instance, amounts to -118.4 W m \(^{-2}\) versus -111.9
W m \(^{-2}\) at a grid spacing of 80 km and 2.5 km, respectively. The changes
in radiative cooling cannot be compensated by changes in sensible heat flux
only and thus generally lead to weaker precipitation amounts at higher res-
olution and, from a water conservation perspective, smaller latent heat flux. In fact, the difference in the radiative cooling of the atmosphere between the 5-km and the 2.5-km simulation, which amounts to -1.3 W m\(^{-2}\), is almost perfectly compensated by changes in sensible heat flux (difference of +0.3 W m\(^{-2}\) between the 5- and 2.5-km simulations) and precipitation (+1.1 W m\(^{-2}\)) or equivalently latent heat flux (+1.0 W m\(^{-2}\)). These relationships become less accurate the coarser the resolution is, with changes in radiative cooling that are larger than changes in precipitation and in latent heat flux, implying changes in the heat storage term.

Although difficult to interpret, it is still interesting to compare the resolution dependencies of the water and energy budgets, as obtained in our ICON simulations that employ explicit convection, to resolution dependencies obtained in a set of GCMs using convective parameterizations and integrated at various resolutions, from about 100 km down to 25 km (see Vanniere et al. 2018). The only two points of agreement concern the sensible heat flux, which also does not vary with resolution in Vanniere et al. (2018) and the outgoing shortwave radiation. The latter, as in the ICON simulations, decreases with increasing resolution, although with a maximum difference of only about 5 W m\(^{-2}\) (see their Figure 4) against 40 W m\(^{-2}\) in our simulations (Figure 1c). For the remaining variables, either they do not exhibit a strong resolution dependency (for net shortwave and net longwave...
radiation at the surface), they do exhibit a strong sensitivity in contrast to ICON (for outgoing longwave radiation) or they exhibit opposite trends (for latent heat flux and precipitation). Also, the resolution dependencies appear to be much weaker in Vanniere et al. (2018).

Figure 4 is a repetition of the analysis of Figure 1 but separated between land and ocean. We only consider the tropical region because, as evident from the previous discussion and from the predominance of convection over the tropics, much of the resolution dependencies arises in the tropics. Like on the global scale and given our metric to assess resolution dependencies, a grid spacing of 5 km appears sufficient to capture the basic statistics of the water and energy budget over land and over ocean. A further refinement of the grid spacing to 2.5 km leads to differences that are smaller than the DYAMOND spread. Breaking up the global response in its land and ocean component reveals a few interesting features. Net shortwave radiation at the surface (Figures 4b,e) and at the top of the atmosphere (Figures 4c,f) as well as surface net longwave radiation (Figures 4b,e) are more sensitive to resolution over tropical ocean than over tropical land. This can be explained by the predominance of shallow clouds over tropical ocean. Also, a similar change in shallow cloud amount over tropical land and over tropical ocean would have a stronger effect on the radiation budget over the darker oceanic surface. In contrast, outgoing longwave radiation (Figure 4c,f), if anything,
rather responds to resolution changes over tropical land.

The increase in surface net shortwave radiation with finer resolution is almost exclusively consumed by a corresponding increase in sensible heat flux over the tropical land area (Figure 4e). The latent heat flux rather decreases with finer resolution, which indicates limitations in the availability of soil moisture to convert the surplus of available energy in latent heating at high resolution. The surface energy budget is almost closed over the tropical land area, with values of 5, 2.9, 1.5, 1.3, 0.9 and 0.8 W m\(^{-2}\) in the 80-, 40-, 20-, 10-, 5- and 2.5-km simulations, respectively. In agreement with this positive energy imbalance, the tropical land surface temperature warms with time, but for the 40 days considered, none of the lower-resolution simulations appears to systematically drift from the 2.5-km simulation. Over tropical ocean (Figure 4b), the use of prescribed SST seems to prevent an increase of both sensible and latent heat fluxes at higher resolutions, as would be expected from the increase in surface net shortwave radiation and the imbalance in the surface energy budget is far from being closed.

Interesting is the distinct precipitation response over tropical land and over tropical ocean (Figures 4a,d). In agreement with studies based on GCMs that have to rely on convective parameterizations (Vanniere et al. 2018; Demory et al. 2013), refining the grid spacing leads to a decrease of precipitation over tropical ocean and to an increase of precipitation over
tropical land. As the latent heat flux actually decreases over land in the tropics (Figure 4d), this implies a stronger transport of moisture from the tropical ocean to the tropical land as well as a stronger transport of moisture from the extratropics to the tropics. The better representation of orography and of land-sea boundaries, in particular over the maritime continent, favors stronger precipitation over land (Qian 2008; Schiemann 2014) and contributes to the observed trend with resolution (see e.g. Fig. 5).

3.2 Tropical climate

Figure 5 shows maps of 40-day mean precipitation for the set of ICON simulations. Qualitatively, large-scale features of the precipitation distribution look very similar across resolutions, in particular concerning the eastern Pacific and the Atlantic ITCZ as well as precipitation over Africa and South America. Larger discrepancies become apparent over the western Pacific and around the maritime continent, an expression of the more complex geographical distribution of the land masses over that region (Qian 2008) and of the spatially more uniform distribution of SSTs. There is also a tendency for larger precipitation maxima to occur as the grid spacing is coarsened, e.g. over the western Pacific and Indian Ocean, as expected from simulations where convection is strongly underresolved and grid-point storms occur (Weisman et al. 1997). Finally, careful inspection of Figure
indicates a less extended second precipitation band over the south western Pacific and a less pronounced separation between this band and the extratropical storm tracks (see around Fiji island) as the grid spacing is refined. This may be interpreted as a less pronounced double ITCZ at finer resolutions.

Figure 5 gives a qualitative overview of the precipitation distribution. To get a more quantitative view, we compute location and width of the Pacific and Atlantic ITCZs (see Figure 6). Those are both first order features that a GCM ought to be able to represent, but that GCMs with convective parameterizations struggle with, with too wide as well as misplaced ITCZs, both over the Pacific and Atlantic (Stanfield et al. 2015; Siongco et al. 2014). Location and width are computed by defining precipitation objects as inspired from the SAL (Structure, Amplitude and Location) measure (Wernli et al. 2008) and for instance applied in Siongco et al. (2014) for evaluating the representation of the Atlantic ITCZ in GCMs. The precipitation objects are defined by the 10 mm day$^{-1}$ precipitation contour. The centroid and minor axis length of the objects, assuming ellipitical objects, give location and meridional width of the ITCZ. Visual inspection of Figure 5 reveals a good agreement between the so derived location and width of the ITCZ and the precipitation distribution. For this analysis, we focus on the eastern Pacific and the Atlantic ITCZs as no robust behavior emerges.
over the western Pacific where the simulations exhibit distinct numbers of precipitation objects.

Figure 6a indicates that the latitudinal position of the Atlantic ITCZ is insensitive to changes in resolution. In contrast, a grid spacing of at least 5 km is required to capture its longitudinal position. Finer grid spacing tends here to shift the Atlantic ITCZ eastward. A similar shift has been observed in GCMs using convective parameterizations (Siongco et al. 2014). In the DYAMOND ensemble, the ITCZ longitudinal position shifts eastwards from ICON (grid spacing of 2.5 km), to IFS (4.8 km), MPAS (3.8 km), NICAM (3.5 km), ARPEGE-NH (2.5 km) and SAM (4.3 km). Except for ICON and SAM, a tendency for the longitudinal position of the ITCZ to shift eastwards with finer grid spacing is thus also apparent in storm-resolving models. The resolution dependency of the position of the Pacific ITCZ is opposite to the one of the Atlantic ITCZ. The longitudinal position of the Pacific ITCZ is insensitive to resolution, whereas capturing its latitudinal position requires a grid spacing of 5 km with a tendency for a more equatorward position at higher resolution (Figure 6c). The contradictory robustness of the latitudinal versus longitudinal position of the ITCZ over the Atlantic and over the Pacific is partly a result of the very large spread in the longitudinal position of the Pacific ITCZ in the DYAMOND ensemble (Figures 6a,c). But it also reflects the fact that, over the Atlantic, coarsening the grid spacing leads to
an increase in precipitation over its western side, whereas over the Pacific, coarsening the grid spacing tends to add precipitation on the north, north-eastern flank of the ITCZ (see Figures 5 and 7). This is consistent with the distinct distribution of SST over the two regions. Over the equatorial Atlantic, the SST gradient is from West to East with a maximum West of the ITCZ location; over the Pacific, between 160° W and 80° W, the region enclosing the precipitation object, the region of maximum SST is located north, north-east from the ITCZ location. Hence, in both cases, coarsening the grid spacing, which makes it more difficult to trigger convection, tends to favor precipitation over the higher SST. The latter is nevertheless positioned differently relative to the precipitation object over the eastern Pacific and over the Atlantic. Conversely, as the finer-resolution simulations can rain both over low and high SSTs and as the total precipitation amount over the tropical oceanic region varies only slightly across resolutions (see e.g. Figure 4a), the finer-resolution simulations will tend to rain more over low SST and less over high SST in comparison to the coarser-resolution simulations.

Concerning the ITCZ width (Figure 6b), changing the resolution has a smaller effect than changing the model formulation over the Atlantic. Even if the resolution-induced differences are smaller than the DYAMOND spread at all resolutions, there is a tendency for a narrowing of the ITCZ
at grid spacings finer than 10 km. Over the eastern Pacific (Figure 6d), the dependency is less systematic with first a broadening, from 80 km down to 20 km, and then a narrowing of the ITCZ. Here the resolution matters and the resolution-induced differences remain always larger than the DYAMOND spread, even at 5 km.

To summarize our findings, Figure 7 attempts to redraw our conceptual picture of the tropical climate based on the results of the simulations for the Pacific ITCZ. A similar picture emerges for the Atlantic ITCZ, though with a much less pronounced increase of precipitation on the northern flank of the ITCZ when the grid spacing is coarsened. For clarity, we only show the 2.5-km and 80-km simulations. We also define on Figure 7 the three cloud categories shallow, congestus and deep based on a cloud top below 1 km, below 4 km and above 8 km, respectively.

As visible on Figure 7, both simulations exhibit a placement of the ITCZ in the northern hemisphere with trade winds converging towards the location of the precipitation maximum at the core of the ITCZ. The trade winds reach values between 6 and 7 m s$^{-1}$ in the subtropics and decelerate to a few m s$^{-1}$ in the doldrums. At the same time, the clouds transition from shallow to deeper clouds, with a successive increase of the fraction of congestus and deep convective clouds in both simulations. Comparing the cloud partitioning between the two simulations, the two simulations
agree on the amount of shallow convection outside the main ITCZ region. They both robustly populate the southern side of the ITCZ with 99% of shallow clouds and mainly shallow clouds on the northern ITCZ flank, with a partitioning of 94% with a grid spacing of 80 km and 88% at 2.5 km. The differences in the amount of congestus clouds simulated by the two simulations across the ITCZ appear also to be small, with values around 30% at the peak precipitation location and around 15% 6° north of it in both simulations. Although a direct comparison to studies that have used satellite observations to quantify the cloud partitioning is difficult due to the wide variety of employed cloud definitions, 30% of all clouds being congestus is larger than what was found in Wall et al. (2013) using CloudSat (see their Fig. 4). In contrast to these robust features, the fraction of deep convective clouds in the main ITCZ region is far from being robust. The values are 9 and 4% at the two considered latitudes for the 80-km simulation versus 34 and 20% for the 2.5-km simulation. These resolution dependencies, as illustrated by Figure 7 for the finest and coarsest simulations, remain valid when considering the range of tested resolutions. The population of shallow convection outside the main ITCZ region and the population of congestus clouds in the ITCZ region remain fairly constant across resolutions with maximum resolution differences below 10%. In contrast, the fraction of deep convective clouds starts to sharply increase by a grid spacing finer
than 20 km at the expense of the shallow cloud population. The values for the fraction of deep convective clouds derived for the locations of the precipitation maximum are 9%, 13%, 14%, 25%, 30% and 34% for a grid spacing of 80 km down to 2.5 km.

3.3 Extratropical climate

One basic and key element of the extratropical circulation is the jet stream. Although not directly linked to cloud and convective processes, GCMs using grid spacing of a few degrees also struggle to capture the location of the jet stream and of the storm tracks with a too equatorward position (Swart and Fyfe 2012). As the jet is more pronounced and more important for the winter climate, we focus our analysis in this subsection on the southern hemisphere. Inspired from our previous analysis on precipitation objects, we diagnose the location of the polar jet in a quantitative way by defining a jet object as a set of connected points where the value of the 100-hPa wind velocity reaches at least 85% of its maximum value. In contrast to the previously defined precipitation objects, we do not use a fixed threshold but a variable threshold to accommodate for the fact that the strength of the polar jet varies by as much as 6 m s$^{-1}$ across the simulations. Also we use 100 hPa (about 15.5 km) rather than a lower pressure level to identify the polar jet as this naturally eliminates the signature of the
subtropical jet. The resulting extent of the jet objects for each simulation as well as their centroid are shown in Figure 8. The latitudinal position of the polar jet shows a weak sensitivity to the grid spacing with values scattered between 51.3° S and 53.9° S and no systematic dependency upon resolution. The longitudinal position of the polar jet also exhibits no systematic dependency but exhibits a much larger variability with values comprised between 22.2° E to 66.4° E. There nevertheless seems to be a tendency for more extended jet objects the coarser the resolution is. These results stand in rough agreement with previous studies that have used GCMs with parameterized convection and other metrics to define the position of the polar jet. Such studies have loosely concluded that by a grid spacing of 1°, the latitudinal position of the polar jet converges towards its observed position (Arakelian and Codron 2012; Pope and Stratton 2002). In a more recent and detailed analysis of the dependency of the polar jet on resolution based on aquaplanet simulations and spanning grid spacings ranging from 300 km down to 28 km, Lu et al. (2015) concluded that the latitudinal position and the intensity of the jet show signs of convergence for grid spacing finer than 50 km.

Fig. 8
4. Discussion

The previous section interpreted resolution differences with respect to the spread of the DYAMOND ensemble. From this we concluded that the resolution differences between the 5-km and 2.5-km simulation are always smaller than the ensemble spread, except for the width of the Pacific ITCZ. This statement is nevertheless conditioned on using the 2.5-km simulation as a reference. Further refining the grid spacing could invalidate this statement for those variables that haven’t converged yet. In a strict sense, we understand under convergence the fact that at a given grid spacing resolution-induced differences become zero. In a more practical sense, the differences will never become zero, but at least they should get smaller with successive grid refinements.

With respect to the components of the water and energy budgets and except for outgoing longwave radiation on the global scale and over tropical ocean, as well as sensible heat flux over tropical land area, none of the variables has reached convergence in a strict sense. This can be recognized by the non-zero differences between the 5-km and 2.5-km simulations in Figs. 1 and 4. However, both for net shortwave radiation as well as surface net longwave radiation, signs of convergence are visible with a continuous reduction of the differences between successive grid refinements (see Figures 1b,c). This is promising as net shortwave radiation and surface net longwave
radiation by far depicted the largest resolution dependencies. In terms of the Atlantic and Pacific ITCZs, their properties do not seem to have converged yet.

The previous section also documented differences between model simulations without any reference to observations. As a last step, we use observations to assess whether refining the grid spacing leads to a better agreement with observations. This allows us to investigate whether variables exhibiting larger resolution dependencies are associated with larger biases.

Mean values derived from observations are listed in Table 2, together with the corresponding values of the 2.5-km and 80-km ICON simulations. Given the fact that most of the differences between the simulations are related to the tropics and that some of the observational datasets do not have global coverage, we restrict our analysis to the tropical area (30° S to 30° N). The resolution sensitivities documented in section 3 and displayed in Figure 1 on the global scale nevertheless remain valid for the tropics as they are mainly the result of convective processes.

For all the considered statistics, refining the grid spacing systematically leads to a better agreement with observations. For instance, precipitation is overestimated at all grid spacings, so the obtained decrease in precipitation with higher resolution brings the simulated values closer to the observed ones. Only for net shortwave radiation and net longwave radiation at the
surface, refining the grid spacing down to 2.5 km ends up to overcorrect
the biases originally present at 80 km. The resulting biases in surface net
shortwave radiation at 2.5 km nevertheless are much smaller than the ones
at 80 km with values of 1 W m$^{-2}$ at 2.5 km versus -56 W m$^{-2}$ at 80 km,
whereas the biases in surface net longwave radiation are more similar, but
of opposite sign, with values of -5.7 W m$^{-2}$ at 2.5 km versus 6.6 W m$^{-2}$ at
80 km (Table 2). It is remarkable that, even for the location and width of
the Atlantic and Pacific ITCZs, as well as for the latitude of the location of
the polar jet, the 2.5-km simulation is the closest to observations.

Our previous analysis also indicated that changes in grid spacing primar-
ily affect the radiation components, except for outgoing longwave radiation.
The biases for the 2.5-km simulation are -9.9 W m$^{-2}$, 1 W m$^{-2}$ and -5.7
W m$^{-2}$ for net shortwave radiation at the top of the atmosphere, at the
surface and for surface net longwave radiation, respectively (Table 2). In
comparison, the biases are -4 W m$^{-2}$ for net longwave radiation at the top
of the atmosphere, 9.5 W m$^{-2}$ for sensible heat flux, 29.5 W m$^{-2}$ for latent
heat flux and 7.8 W m$^{-2}$ for precipitation (Table 2). It follows that small
resolution dependencies do not say much about the quality of a simulation.
Some model deficiencies are not improved by resolution and changes in the
model physics, dynamics or in the model parameter settings would have a
larger impact.
5. Conclusions

We examined the resolution dependencies of global simulations conducted with the ICON model and using grid spacings of 80, 40, 20, 10, 5 and 2.5 km. The focus was on the behavior of basic climate statistics, namely water and energy budgets, location and width of the ITCZ, trimodal cloud distribution in the tropics, as well as jet positions in the extratropics. Those are all basic statistics and dominating features of the large-scale circulation that a GCM ought to be able to correctly represent. All the simulations used explicit convection, the same physical parameterizations and were not retuned for a specific grid spacing. To objectively quantify the resolution dependencies, we compared resolution-induced differences to the spread obtained in the DYAMOND ensemble of models. The latter comprises eight global climate models integrated at storm-resolving resolution of O(3 km) for 40 days. Despite the shortness of the simulation period, we believe that the results could hold to longer time period as differences developed within a few days and did not appear to drift with time. The main results are as follows:

• For all the considered twenty-seven statistics and with the exception of the width of the Pacific ITCZ, at least a grid spacing of 5 km appears sufficient to capture the basic properties of the climate system.
Further refining the grid spacing to 2.5 km leads to differences that are smaller than changing the model physics/dynamics as measured by the DYAMOND ensemble spread. For several of the considered statistics, namely global mean precipitation, global mean sensible heat flux, global mean outgoing longwave radiation, mean tropical oceanic sensible heat flux, mean tropical oceanic outgoing longwave radiation, mean precipitation over tropical land area, mean latent heat flux over tropical land area, latitudinal position and width of the Atlantic ITCZ as well as longitudinal position of the Pacific ITCZ, even coarsening the grid spacing to 80 km leads to differences that are smaller than the DYAMOND ensemble spread.

- The largest resolution-induced differences are found in net shortwave radiation, both at the surface and at the top of the atmosphere, with differences of up to about 40 W m\(^{-2}\), and to a lesser extent in surface net longwave radiation (10 W m\(^{-2}\)). The differences result from a systematic deepening of the planetary boundary layer and an increase in low clouds over the subtropical oceans as the grid spacing is coarsened. This leads to an increase in net shortwave radiation, both at the surface and at the top-of-the atmosphere, with finer grid spacing and to an increase in the surface downward longwave radiation.
Resolution only matters for the longitudinal position of the Atlantic ITCZ, whereas for the eastern Pacific ITCZ, it only affects its latitudinal position. The Atlantic ITCZ tends to shift eastward with finer grid spacing, whereas the Pacific ITCZ tends to shift equatorward. This distinct behavior can be related to the distinct orientation of the SST gradient over both regions.

All the simulations exhibit the expected deepening of convective clouds in the ITCZ. The congestus cloud population makes 20 to 29% of all clouds at the location of peak precipitation depending upon the grid spacing. In contrast, the deep convective cloud population systematically increases at the expanse of the shallow cloud population when refining the grid spacing. The fraction of deep convective clouds varies from 9% (at 80 km) to 34% (at 2.5 km).

In the extratropics, the jet position shows no systematic resolution dependency but its longitudinal position exhibits strong resolution dependencies.

Refining the grid spacing ends up always reducing simulation biases, but statistics exhibiting less resolution dependency are not necessarily closer to observations.

Even if resolution differences become smaller than the DYAMOND en-
semble spread by a grid spacing of 5 km at least, the resolution differences between the 5- and 2.5-km simulations are not zero yet, meaning that the investigated climate statistics haven’t converged in a strict sense. However, net shortwave radiation, which depicts the largest resolution dependency, shows sign of convergence with resolution differences getting smaller with successive grid refinements. Combined with the overall small resolution dependencies as assessed by the comparison to the DYAMOND ensemble spread, we conclude that simulations with a grid spacing of 5 km and using explicit convection may be used to simulate the climate. Even a grid spacing of 10 km would be sufficient to capture eighteen out of the twenty-seven investigated statistics. This is promising given the computational burden associated with simulations using a grid spacing of 2.5 km and seems to confirm experience with the NICAM model. Only for the width of the Pacific ITCZ and for the fraction of deep convective clouds, large resolution dependencies remain that could question the use of a grid spacing of 5 km.

Acknowledgements

The authors thank the ESiWACE project for technical support as well as provision of computing time and storage space. DK acknowledges funding from the Hans Ertel Center for Weather Research (HErZ). This German
research network of universities, research institutions and DWD is funded by the BMVI (Federal Ministry of Transport and Digital Infrastructure). GC acknowledges funding from the HD(CP)$^2$ project. Use of the supercomputer facilities at the Deutsches Klimarechenzentrum (DKRZ), technical support of Leonidas Linardakis, comments by Hauke Schmidt on an earlier version of this manuscript, as well as comments of the reviewers are acknowledged. Primary data and scripts used in the analysis and other supplementary information that may be useful in reproducing the authors’ work are archived by the Max Planck Institute for Meteorology and can be obtained by contacting publications@mpimet.mpg.de.

References


Baldauf, M., A. Seifert, J. Förstner, D. Majewski, M. Raschendorfer, and


Yashiro, H., Y. Kajikawa, Y. Miyamoto, R. Yoshida, T. Yamaura, and H. Tomita, 2016: Resolution dependency of diurnal precipitation

List of Figures

1. Global statistics from the ICON simulations, with a grid spacing of 80 km (on the left of the x-axis) down to 2.5 km (on the right of the x-axis), averaged over the 40-day period and expressed as difference to the 2.5-km simulation: (a) precipitation, (b) surface components of the energy budget and (c) top-of-the-atmosphere (TOA) components of the energy budget. The vertical bars show one standard deviation of the corresponding quantities in the DYAMOND ensemble. Downward energy fluxes are taken as positive except for sensible and latent heat fluxes which are taken as positive when directed into the atmosphere. For precipitation, the GEOS model is not included in the computation of the DYAMOND spread as there were errors in the precipitation diagnostic.

2. Difference map between the 80-km and 2.5-km simulations for (a) net shortwave radiation at the surface and (b) cloud cover. Note that the data have been regridded on a 0.75° grid (approximately the grid spacing of the 80-km simulation) and averaged over the 40-day period.

3. Profiles of (a) temperature difference, (b) specific humidity and (c) diagnostic cloud water after 12 hours of simulation averaged over the south-eastern Pacific (30°S-0°; 140°W-85°W), one of the region exhibiting large differences in cloud cover as an example. Panel (a) shows difference between the 80- and 2.5-km simulation, whereas panels (b-c) show the 80-km simulation in red and the 2.5-km simulation in black. The diagnostic cloud water includes contributions from the subgrid cloud cover scheme.

4. Same as Figure 1 but averaged over (a-c) tropical ocean and (d-f) tropical land. Tropics are from 30°S to 30°N.
Tropical mean precipitation (mm day$^{-1}$) as a function of grid spacing: (a) 80 km, (b) 40 km, (c) 20 km, (d) 10 km, (e) 5 km and (f) 2.5 km. Contour (orange line), location (dot) and width (red line) of the precipitation objects used to define the Pacific and Atlantic ITCZs are overlaid. The data have been regridded on a 0.75° grid (approximately the grid spacing of the 80-km simulation) to allow a quantitative comparison of the precipitation objects.

(a,c) Location of the ITCZ over the Atlantic and Pacific as well as (b,d) corresponding width. Grey bar for the spread of the DYAMOND ensemble (computed as one standard deviation). The GEOS model is not included in the computation of the spread as there were errors in the precipitation diagnostic. Note that all the data have been regridded on a 0.75° grid (approximately the grid spacing of the 80-km simulation) for fair comparison.

Picture of the ITCZ with precipitation (mm day$^{-1}$, line), wind velocity (m s$^{-1}$, arrow) and cloud distribution (%, number) derived from the results of the 2.5-km (black) and 80-km (red) simulation for the Pacific ITCZ. Only the eastern side of the Pacific, comprised between 170°-90° W (see location of the precipitation object in Figure 5) is considered. The precipitation curves correspond to the zonally averaged precipitation values, whereas the wind arrows indicate the zonal average of the wind velocity plotted every 6° from the location of the maximum precipitation in each simulation. The cloud categories are defined based on cloud top, defined as the first height where the sum of cloud liquid water content and cloud ice drops below 10$^{-3}$ g kg$^{-1}$. Only clouds with a base below 1 km are considered. Shallow clouds (SCu) have the top below 4 km, congestus clouds (Cog) between 4 and 8 km, and deep clouds (DCu) above 8 km. The cloud partitioning is computed as a function of latitude and plotted every 6° from the location of the maximum precipitation in each simulation.

Location (crosses) and extent (contour lines) of the jet objects for the set of simulations. See text for more detail.
Fig. 1. Global statistics from the ICON simulations, with a grid spacing of 80 km (on the left of the x-axis) down to 2.5 km (on the right of the x-axis), averaged over the 40-day period and expressed as difference to the 2.5-km simulation: (a) precipitation, (b) surface components of the energy budget and (c) top-of-the atmosphere (TOA) components of the energy budget. The vertical bars show one standard deviation of the corresponding quantities in the DYAMOND ensemble. Downward energy fluxes are taken as positive except for sensible and latent heat fluxes which are taken as positive when directed into the atmosphere. For precipitation, the GEOS model is not included in the computation of the DYAMOND spread as there were errors in the precipitation diagnostic.
Fig. 2. Difference map between the 80-km and 2.5-km simulations for (a) net shortwave radiation at the surface and (b) cloud cover. Note that the data have been regridded on a 0.75° grid (approximately the grid spacing of the 80-km simulation) and averaged over the 40-day period.
Fig. 3. Profiles of (a) temperature difference, (b) specific humidity and (c) diagnostic cloud water after 12 hours of simulation averaged over the south-eastern Pacific (30°S-0°; 140°W-85°W), one of the region exhibiting large differences in cloud cover as an example. Panel (a) shows difference between the 80- and 2.5-km simulation, whereas panels (b-c) show the 80-km simulation in red and the 2.5-km simulation in black. The diagnostic cloud water includes contributions from the subgrid cloud cover scheme.
Fig. 4. Same as Figure 1 but averaged over (a-c) tropical ocean and (d-f) tropical land. Tropics are from 30°S to 30°N.
Fig. 5. Tropical mean precipitation (mm day$^{-1}$) as a function of grid spacing: (a) 80 km, (b) 40 km, (c) 20 km, (d) 10 km, (e) 5 km and (f) 2.5 km. Contour (orange line), location (dot) and width (red line) of the precipitation objects used to define the Pacific and Atlantic ITCZs are overlaid. The data have been regridded on a 0.75° grid (approximately the grid spacing of the 80-km simulation) to allow a quantitative comparison of the precipitation objects.
Fig. 6. (a,c) Location of the ITCZ over the Atlantic and Pacific as well as (b,d) corresponding width. Grey bar for the spread of the DYAMOND ensemble (computed as one standard deviation). The GEOS model is not included in the computation of the spread as there were errors in the precipitation diagnostic. Note that all the data have been regridded on a 0.75° grid (approximately the grid spacing of the 80-km simulation) for fair comparison.
Fig. 7. Picture of the ITCZ with precipitation (mm day$^{-1}$, line), wind velocity (m s$^{-1}$, arrow) and cloud distribution (%, number) derived from the results of the 2.5-km (black) and 80-km (red) simulation for the Pacific ITCZ. Only the eastern side of the Pacific, comprised between 170°-90° W (see location of the precipitation object in Figure 5) is considered. The precipitation curves correspond to the zonally averaged precipitation values, whereas the wind arrows indicate the zonal average of the wind velocity plotted every 6° from the location of the maximum precipitation in each simulation. The cloud categories are defined based on cloud top, defined as the first height where the sum of cloud liquid water content and cloud ice drops below $10^{-3}$ g kg$^{-1}$. Only clouds with a base below 1 km are considered. Shallow clouds (SCu) have the top below 4 km, congestus clouds (Cog) between 4 and 8 km, and deep clouds (DCu) above 8 km. The cloud partitioning is computed as a function of latitude and plotted every 6° from the location of the maximum precipitation in each simulation.
Fig. 8. Location (crosses) and extent (contour lines) of the jet objects for the set of simulations. See text for more detail.
1 Experimental set-up and model performance. The grid spacing is computed as the square root of the mean cell area of the model triangles. Each node contains 24 cores and the maximum memory is per node. Number in parenthesis by the wall clock time indicates the expected wall clock time scaling up from the previous resolution taking into account differences in grid spacing, time step and number of nodes. Likewise for the maximum memory.  

2 Various statistics averaged over the simulation period for the 80- and 2.5-km simulations as well as for observations. Note that for latent and sensible heat fluxes, observations are only available over ocean so the simulated tropical values are for ocean only (whereas the global values include land and ocean). Units are mm day$^{-1}$ for precipitation, W m$^{-2}$ for the energy fluxes, and ° for the Atlantic ITCZ, Pacific ITCZ and the polar jet. Observations are taken from the TRMM Multi-satellite Precipitation Algorithm for precipitation (including the ITCZs statistics), from the CERES syn1deg product for the surface radiative fluxes, from the CERES ssf1deg product for the top-of-the-atmosphere fluxes and from the OAFlux for the surface fluxes. For the center of the polar jet, ERA-Interim is used. All data were regridded to the same 0.75°x0.75° grid spacing. Tropics are from 30° S to 30° N.
Table 1. Experimental set-up and model performance. The grid spacing is computed as the square root of the mean cell area of the model triangles. Each node contains 24 cores and the maximum memory is per node. Number in parenthesis by the wall clock time indicates the expected wall clock time scaling up from the previous resolution taking into account differences in grid spacing, time step and number of nodes. Likewise for the maximum memory.

<table>
<thead>
<tr>
<th>Resolution</th>
<th>Grid spacing (km)</th>
<th>Time step (s)</th>
<th>Nodes</th>
<th>Wall clock time (min)</th>
<th>Max memory (MB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R2B5</td>
<td>80</td>
<td>450</td>
<td>24</td>
<td>33</td>
<td>400</td>
</tr>
<tr>
<td>R2B6</td>
<td>40</td>
<td>225</td>
<td>96</td>
<td>51 (66)</td>
<td>430 (400)</td>
</tr>
<tr>
<td>R2B7</td>
<td>20</td>
<td>150</td>
<td>96</td>
<td>209 (306)</td>
<td>840 (1720)</td>
</tr>
<tr>
<td>R2B8</td>
<td>10</td>
<td>90</td>
<td>420</td>
<td>380 (318)</td>
<td>1000 (768)</td>
</tr>
<tr>
<td>R2B9</td>
<td>5</td>
<td>45</td>
<td>420</td>
<td>1824 (3040)</td>
<td>2600 (4000)</td>
</tr>
<tr>
<td>R2B10</td>
<td>2.5</td>
<td>22.5</td>
<td>540</td>
<td>10320 (11349)</td>
<td>9500 (8089)</td>
</tr>
</tbody>
</table>
Table 2. Various statistics averaged over the simulation period for the 80- and 2.5-km simulations as well as for observations. Note that for latent and sensible heat fluxes, observations are only available over ocean so the simulated tropical values are for ocean only (whereas the global values include land and ocean). Units are mm day$^{-1}$ for precipitation, W m$^{-2}$ for the energy fluxes, and ° for the Atlantic ITCZ, Pacific ITCZ and the polar jet. Observations are taken from the TRMM Multisatellite Precipitation Algorithm for precipitation (including the ITCZs statistics), from the CERES syn1deg product for the surface radiative fluxes, from the CERES ssf1deg product for the top-of-the-atmosphere fluxes and from the OAFlux for the surface fluxes. For the center of the polar jet, ERA-Interim is used. All data were regridded to the same 0.75°x0.75° grid spacing. Tropics are from 30° S to 30° N.

<table>
<thead>
<tr>
<th>Variable</th>
<th>ICON 80 km</th>
<th>ICON 2.5 km</th>
<th>Observation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Global</td>
<td>Tropics</td>
<td>Global</td>
</tr>
<tr>
<td>Precipitation</td>
<td>3.25</td>
<td>3.90</td>
<td>3.12</td>
</tr>
<tr>
<td>Latent heat flux</td>
<td>91.8</td>
<td>157.0</td>
<td>89.3</td>
</tr>
<tr>
<td>Sensible heat flux</td>
<td>28.1</td>
<td>20.4</td>
<td>29.8</td>
</tr>
<tr>
<td>Surface net longwave radiation</td>
<td>-47.0</td>
<td>-47.1</td>
<td>-56.9</td>
</tr>
<tr>
<td>Surface net shortwave radiation</td>
<td>124.2</td>
<td>149.0</td>
<td>162.4</td>
</tr>
<tr>
<td>TOA net longwave radiation</td>
<td>-245.3</td>
<td>-265.4</td>
<td>-246.4</td>
</tr>
<tr>
<td>TOA net shortwave radiation</td>
<td>204.2</td>
<td>246.3</td>
<td>240.</td>
</tr>
<tr>
<td>Atlantic ITCZ center</td>
<td>8.80, -31.4</td>
<td>8.37, -28.0</td>
<td>8.38, -24.4</td>
</tr>
<tr>
<td>Pacific ITCZ center</td>
<td>11.0, -122.</td>
<td>9.8, -119.8</td>
<td>10.1, -117.2</td>
</tr>
<tr>
<td>Atlantic ITCZ width</td>
<td>10.0</td>
<td>9.16</td>
<td>8.21</td>
</tr>
<tr>
<td>Pacific ITCZ width</td>
<td>13.9</td>
<td>13.05</td>
<td>12.3</td>
</tr>
<tr>
<td>Polar jet center</td>
<td>-53.9, 22.2</td>
<td>-52.2, 54.7</td>
<td>-52.6, 41.4</td>
</tr>
</tbody>
</table>