Climate Statistics in Global Simulations of the Atmosphere, from 80 to 2.5 km Grid Spacing

Cathy HOHENEGGER, Luis KORNBLUEH

Max Planck Institute for Meteorology, Germany

Daniel KLOCKE

Hans Ertel Centre for Weather Research, Deutscher Wetterdienst, Germany

Tobias BECKER, Guido CIONI

Max Planck Institute for Meteorology, Germany

Jan Frederik ENGELS

Deutsches Klimarechenzentrum, Germany

Uwe SCHULZWEIDA and Bjorn STEVENS

Max Planck Institute for Meteorology, Germany

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 nine (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 the resolution because of reductions in the 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. In addition, in the tropics, the deep convective cloud population systematically increases at the expense 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.

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

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 using a resolution fine enough to explicitly represent the major modes of heat transfer challenging. Biases resulting from the use of parameterizations, particularly convective parameterizations, have been persistent (Flato et al. 2013). Many of the basic properties of the climate system that a GCM ought to be able to represent, such as the width and location of the Intertropical Convergence Zone (ITCZ), water and energy budgets, surface temperature, land sea contrast, and the position of storm tracks, often exhibit large biases; if not, it is often so because they are a target of the tuning procedure (Hourdin et al. 2017). Refining the horizontal grid spacing down to a few kilometers and explicitly representing convection, as in the so-called convection-permitting models, can overcome at least the problem of having to parameterize deep convection. Nevertheless, it remains unclear which horizontal resolution is needed to capture the basic properties of the climate system.

Using 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 modeling (Prein et al. 2015). Turning off the deep convection scheme, despite not being 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 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 owing to changes in their cloud morphology, leading to larger net shortwave radiation at the surface as compared to their parameterized counterparts (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 to be necessary to represent non-hydrostatic dynamics and to avoid grid-point storms. Whereas the properties of individual convective updrafts, such as the area and velocity, require a finer grid spacing than 4 km (Petch et al. 2002; Bryan et al. 2003; Craig and Dörnbrack 2008; Jeevanjee 2017; Hanley et al. 2015), bulk properties, such as the 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, the properties of individual convective updrafts do not always project strongly on the skill of a simulation even if these properties have not converged yet (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 a convection-permitting resolution on a global scale, with grid spacings ranging from 14 km down to 0.87 km. Finer grid spacings lead to later triggering of convection and to a later peak of precipitation, in better agreement with the observations (Sato et al. 2008, 2009; Noda et al. 2012). Yashiro et al. (2016) concluded that a minimum grid spacing of 2–3 km is necessary to capture the main characteristics of the diurnal cycle. Likewise, contingent upon the chosen definition of deep convective cores, the properties of the simulated convective cores were found to

qualitatively change between grid spacings of 3.5 km 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 the 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 that are often tuned toward observations in state-of-the-art GCMs. Those concern water and energy budgets, location and width of the ITCZ, cloud distribution in the tropics, and jet positions in the extratropics. Moreover, the ability to distinguish between land and ocean will be considered. In order 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 intercomparison project of global storm-resolving models, currently comprising nine models and integrated at a grid spacing of O(3 km). The motivation behind using the DYAMOND ensemble as a reference dataset is twofold. First, it allows the objective quantification of the resolution dependency, whereas traditional convergence studies have to, often subjectively, decide when differences 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 (ICON) model (Zängl et al. 2015), from 80 km down to 2.5 km, following the DYAMOND experimental protocol. Changes in grid spacing imply changes in the time step and resolution of the external parameter fields, but parameterizations 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 to be counter-intuitive, past studies (Webb et al. 2015; Maher et al. 2018) have shown that coarse-resolution models can run stably without using a convection scheme and produce a precipitation climatology that captures the observed large-scale features. We also only consider the effect of the horizontal resolution, not the 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 experimental 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 (SST) and sea ice concentration reading. Physical parameterizations in ICON version 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 Rapid Radiative Transfer Model (RRTM) 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).

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. The number in parentheses by the wall clock time indicates the expected wall clock time scaling up from the previous resolution, taking into account differences in the grid spacing, time step, and number of nodes. The same is true for the maximum memory.

Resolution	Grid spacing (km)	Time step (s)	Nodes	Wall clock time (min)	Max memory (MB)
R2B5	80	450	24	33	400
R2B6	40	225	96	51 (66)	430 (400)
R2B7	20	150	96	209 (306)	840 (1720)
R2B8	10	90	420	380 (318)	1000 (768)
R2B9	5	45	420	1824 (3040)	2600 (4000)
R2B10	2.5	22.5	540	10320 (11349)	9500 (8089)

A similarly configured ICON model version has also been recently used for limited-area simulations over the tropical Atlantic by Klocke et al. (2017).

2.2   Experimental set-up

In order to investigate horizontal resolution dependencies, we analyze a set of simulations in which the horizontal grid spacing is successively refined from 80 km down to 2.5 km by a factor 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 employs a convective parameterization, neither for shallow nor for deep convection. 90 levels are used in the vertical with the model top at 75 km. Damping starts in the 77th 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 initialized 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. The daily observed SST 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 simulations with a grid spacing of 5 km and 10 km. For the remaining coarser-resolution simulations, the time step had to be increased by less than the 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, 2 in Asensio et al. 2019).

2.3   Computational aspects

The simulations have been integrated on the supercomputer Mistral of the Deutsches KlimaRechen-Zentrum (DKRZ). The computation partitions comprise 3,300 dual-socket Intel-CPU nodes with around 265 terabytes of main memory. Some numbers on the model's performance are listed in Table 1. They show that, for the chosen number of nodes, ICON scales well with the increases in wall clock time between subsequent grid spacing refinements being less than what would be expected from the increases in the 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 a grid spacing of O(100 m) over Germany. This effort ensured that ICON scales well (Heinze et al. 2017). The most challenging simulation with a grid spacing of 2.5 km could perform six simulated days per day on 540 nodes and is also com-

parable to what was achieved in DYAMOND by other models (Stevens et al. 2019).

The major problems for the finest-resolution simulation were not per se the computing time, but the generation of input data. As we came 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 the discussion in Stevens et al. 2019).

2.4   Analysis methodology

In order to objectively quantify resolution dependencies, we will compare the resolution differences obtained in our ICON simulations to the spread derived from the DYAMOND ensemble of storm-resolving models. Note that, in contrast to the literature, we use here the term "storm-resolving" and not "convection-permitting" to characterize atmospheric models employing grid spacings of a few kilometers (for a discussion of nomenclature, see Satoh et al. 2019). This is 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 does not use a convective parameterization but still produces convection in the tropics. Unless indicated otherwise, the following models of the DYAMOND ensemble are included to compute the spread (see Stevens et al. 2019): ARPEGE-NH (grid spacing: 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 km 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 a 2.5-km grid spacing. This approach will allow us to determine which grid spacing, between 80 km and 2.5 km, is needed to capture the basic statistics of the climate system.

This last statement is conditioned on the way we assess resolution differences. As in studies that have looked into 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, particularly for climate statistics that exhibit strong resolution dependencies and that have not converged yet. This point is investigated in more detail 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 on 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 the 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 the obtained resolution dependencies may also be larger than expected. In this sense, these two effects can partly cancel each other 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 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 reso-
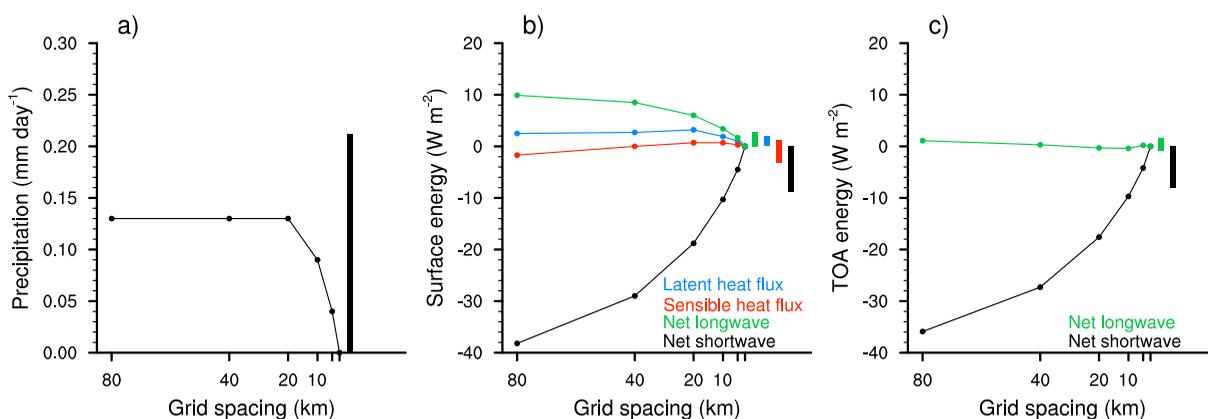
Fig. 1.   Global statistics from the ICON simulations, with a grid spacing of 80 km (on the left of the x-axis) down to 2.5 km (on the right of the x-axis), averaged over the 40-day period and expressed as the difference to the 2.5-km simulation: (a) precipitation, (b) surface components of the energy 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 was not included in the computation of the DYAMOND spread as there were errors in the precipitation diagnostics.

lution 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 periods.

3.   Results

3.1   Water and energy budgets

Figure 1 shows the mean values of the components of the water and energy budgets at the surface and at the top of the atmosphere as a function of the grid spacing of the ICON simulations and expressed as a 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 Fig. 1.

Given our metric to assess resolution differences, Fig. 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, as well as, to a lesser extent, the net long-wave radiation at the surface. Refining the grid spac-

ing from 80 km down to 2.5 km leads to a systematic increase in the net shortwave radiation of about 40 W m^{-2} (Figs. 1b, c), whereas the surface net longwave radiation, with a maximum difference of 10 W m^{-2}, gets systematically more negative (Fig. 1b). Splitting the surface net longwave radiation into 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 the surface net longwave radiation stem 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 SST.

In terms of spatial distribution, the largest differences in the surface net shortwave radiation can be found over the southern tropical Atlantic, the southeastern tropical Pacific, the southern Indian Ocean east of Madagascar reaching almost over to Australia, and over the northeastern subtropical Pacific, as shown in Fig. 2a using the difference between the coarsest and finest resolution as a representative example. These various regions are all regions that also exhibit large

a)

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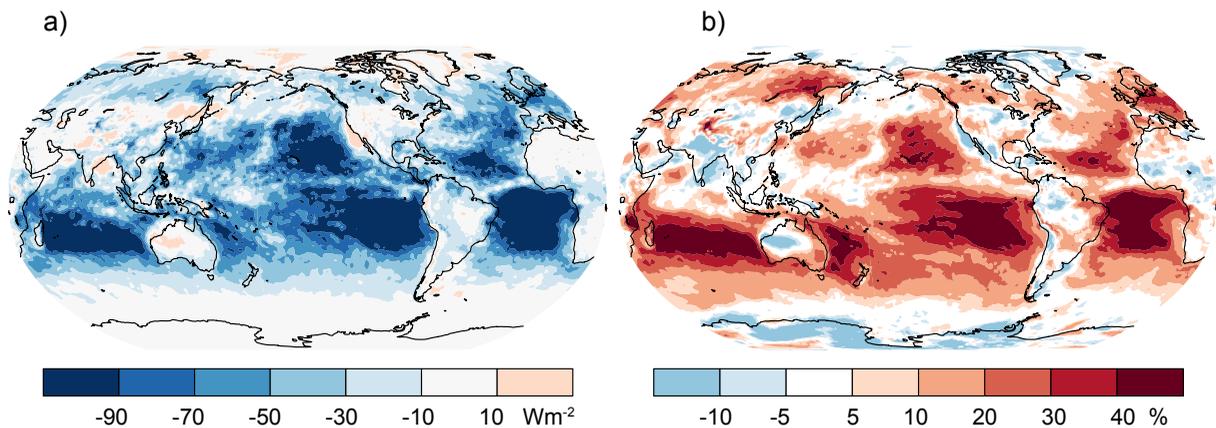
Fig. 2.   Difference map between the 80-km and the 2.5-km simulations for the (a) net shortwave radiation at the surface and (b) cloud cover. Note that the data have been regridded on a 0.75° grid (approximately the grid spacing of the 80-km simulation) and averaged over the 40-day period.

differences in cloud cover (Fig. 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 a much higher cloud amount compared to the 2.5-km simulation (see Fig. 2b). The tendency for the coarser-resolution simulation to produce more cloud cover over the subtropical oceanic region than the finer-resolution simulation remains true when considering the range of the investigated resolutions. The cloud cover averaged from 20°S to 0°S over oceanic areas only, which encompass 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 spacings, which is consistent with and explains the systematic increases in the net shortwave radiation at the surface and at the top of the atmosphere, as well as the decrease in the downward longwave radiation at the surface, as shown in Fig. 1. The cloud liquid water content exhibits a similar behavior over those regions. The decrease in the surface downward longwave radiation with finer grid spacings 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 drying and others exhibiting moistening.

The large differences in cloud cover across the simulations already settle in after 12 h of simulation.

In order to better understand these differences, we look at the profiles of temperature, specific humidity, and cloud water for the two extreme simulations under such conditions (Fig. 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 (500–2,000 m), and the trade inversion around 2 km. The 80-km simulation is associated with a deeper mixed layer (see Fig. 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 (Figs. 3a, c). As a consequence, the cloud layer is more likely to saturate, leading to widespread cloud formation at 80 km (Fig. 3c). The deeper mixed layer at 80 km is consistent with weaker mass transport by shallow cumuli (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 km to 14 km. They concluded that the subgrid-scale cloud parametrization strongly controlled the low-cloud amount and led to an 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 surfaces is small. This is consistent with the negligible sensitivity of outgoing longwave radiation to resolution, as shown in Fig. 1c.
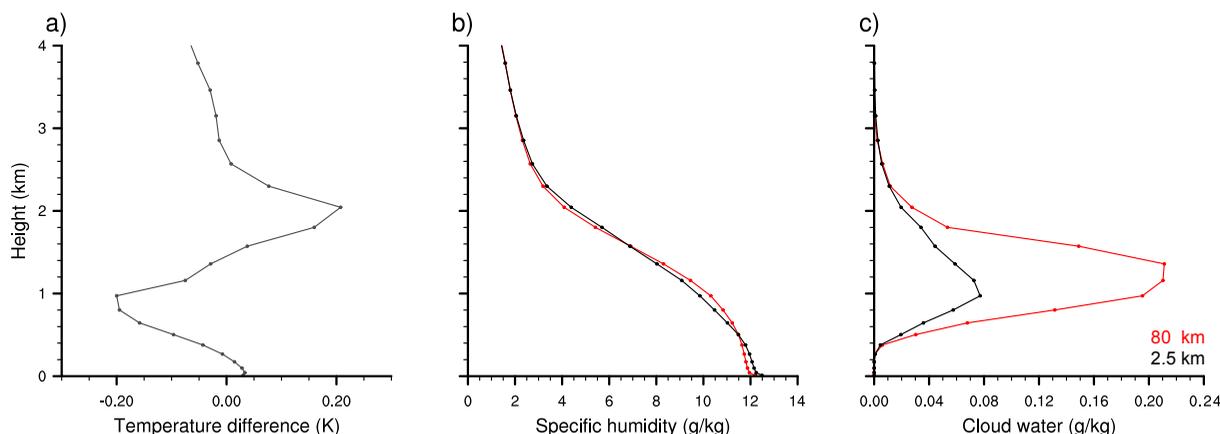
Fig. 3.   Profiles of (a) temperature difference, (b) specific humidity, and (c) diagnostic cloud water after 12 h of simulation averaged over the southeastern Pacific (30°S–0°, 140–85°W), one of the regions exhibiting large differences in cloud cover as an example. Panel (a) shows the difference between the 80-km and the 2.5-km simulations, whereas panels (b) and (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.

Moreover, although simulations with higher horizontal resolutions have less cloud cover in the subtropics and, hence, larger clear-sky areas, 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 gets less negative with finer grid spacings. The values are -247, -248, -246, -243, -241, and -238 W m^{-2} at 80, 40, 20, 10, 5, and 2.5 km, respectively.

The more modest resolution dependencies of precipitation, sensible heat flux, and latent heat flux, as compared to the radiative fluxes (Figs. 1a, b), can be mostly understood from energy constraints. Changes in the net shortwave radiation at the surface and at the top of the atmosphere almost compensate (compare Figs. 1b, c). Given the decrease in the downward longwave radiation at the surface at finer resolutions, the atmospheric radiative cooling, which needs to be balanced by the surface fluxes, decreases. 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. Changes in radiative cooling cannot be compensated by changes in sensible heat flux only and, thus, generally lead to weaker precipitation amounts at higher resolutions 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 simulations, which amounts to -1.3 W m^{-2}, is almost perfectly compensated by changes in the sensible heat flux (difference of +0.3 W m^{-2} between the 5- and the 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 with coarser resolution, with changes in radiative cooling that are larger than changes in precipitation and in latent heat flux, implying changes in the heat storage term.

Despite being 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. 2019). The only two points of agreement concern the sensible heat flux, which also does not vary with resolution in Vanniere et al. (2019), and the outgoing shortwave radiation. The latter, as in the ICON simulations, decreases with increasing resolution, albeit with a maximum difference of only about 5 W m^{-2} (see their Fig. 4) against 40 W m^{-2} in our simulations (Fig. 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). In addition, the resolution dependencies appear to be much weaker in the study of Vanniere et al. (2019).

Figure 4 is a repetition of the analysis in Fig. 1 but
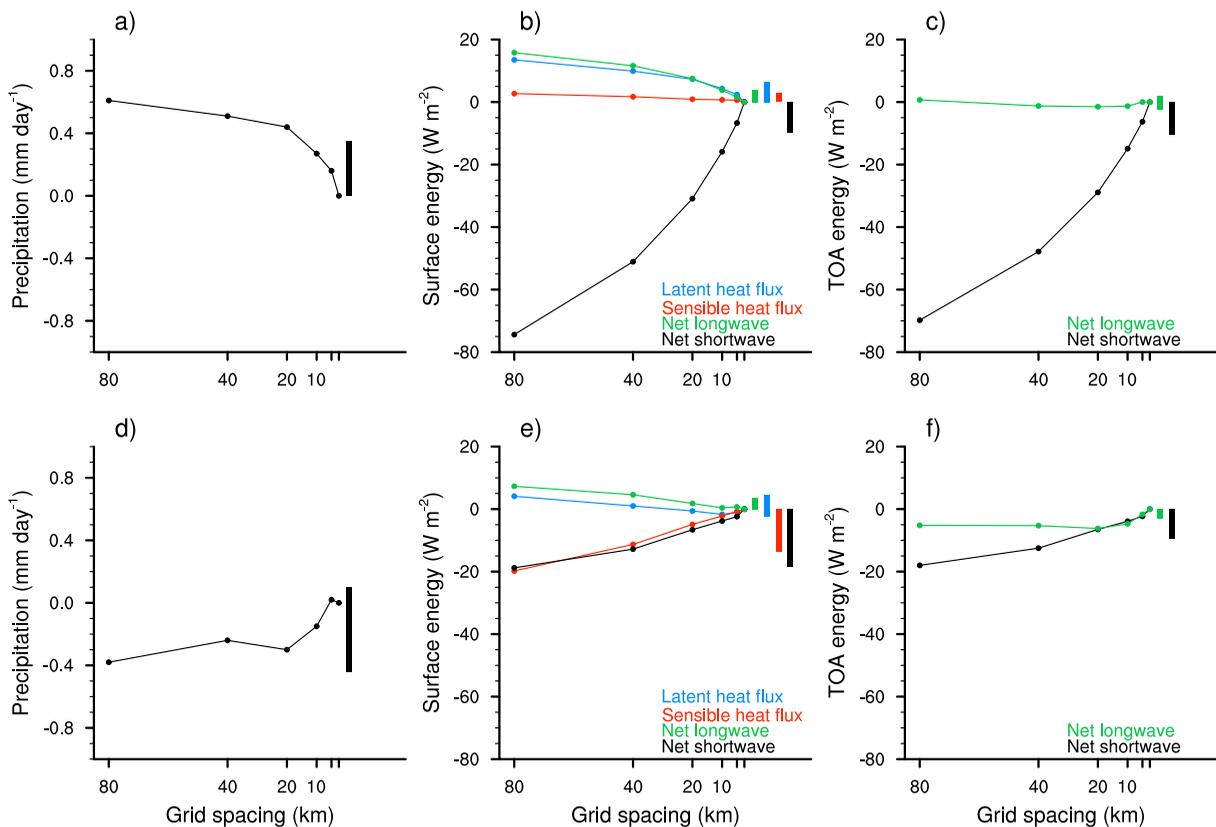
Fig. 4.   Same as Fig. 1 but averaged over (a–c) tropical ocean and (d–f) tropical land. The tropics are from 30°S to 30°N.

with a separation 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 dependency arises in the tropics. Like on the global scale and given our metric to assess resolution dependencies, a grid spacing of 5 km appears to be sufficient to capture the basic statistics of the water and energy budgets 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 into its land and ocean components reveals a few interesting features. The net shortwave radiation at the surface (Figs. 4b, e) and at the top of the atmosphere (Figs. 4c, f) as well as the surface net longwave radiation (Figs. 4b, e) are more sensitive to the resolution over tropical ocean than over tropical land. This can be explained by the predominance of shallow clouds over tropical ocean. In addition, a similar change in the 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, the outgoing longwave radiation (Figs. 4c, f), if anything, rather responds to resolution changes over tropical land.

Over the tropical land area, the increase in the surface net shortwave radiation with finer resolutions is almost exclusively consumed by a corresponding increase in sensible heat flux (Fig. 4e). The latent heat flux rather decreases with finer resolutions, which indicates limitations in the availability of soil moisture to convert the surplus of available energy in latent heating at a 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 rises with time; however, for the 40 days considered, none of the lower-resolution simulations appears to systematically drift from the 2.5-km simulation. Over tropi-
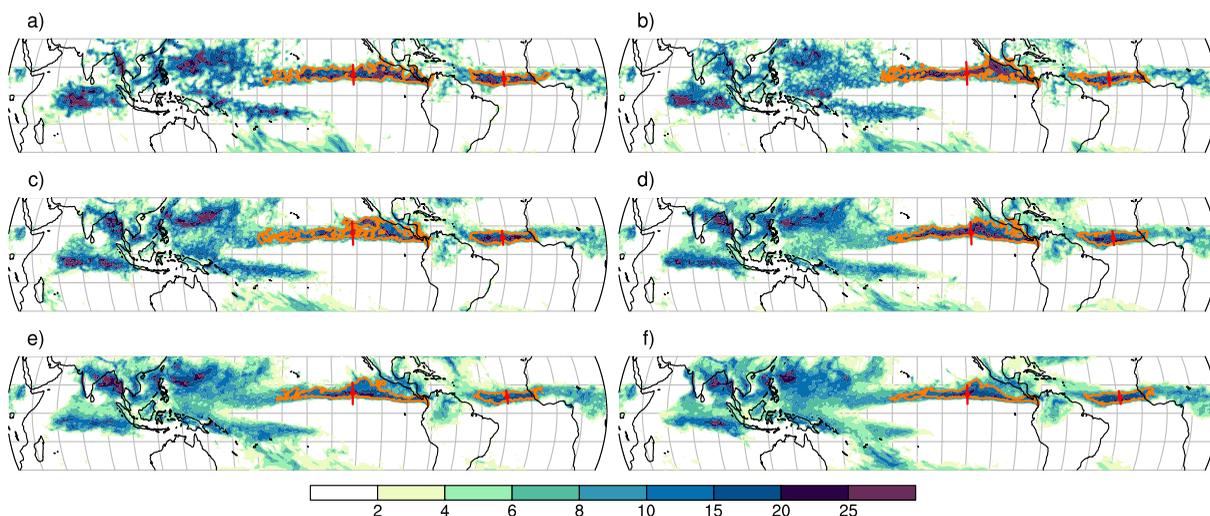
Fig. 5. Tropical mean precipitation (mm day^{-1}) as a function of grid spacing: (a) 80 km, (b) 40 km, (c) 20 km, (d) 10 km, (e) 5 km, and (f) 2.5 km. Contour (orange line), location (dot), and width (red line) of the precipitation objects used to define the Pacific and Atlantic ITCZs are overlaid. The 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.

cal ocean (Fig. 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 the surface net shortwave radiation, and the imbalance in the surface energy budget is far from being closed.

What is interesting is the distinct precipitation response over tropical land and over tropical ocean (Figs. 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 in precipitation over tropical ocean and to an increase in precipitation over tropical land. As the latent heat flux actually decreases over land in the tropics (Fig. 4e), 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, the large-scale features of the precipitation distribution look very similar across resolutions, in particular con-

cerning the eastern Pacific and the Atlantic ITCZs, 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, for example, over the western Pacific and Indian Ocean, as expected from simulations in which convection is strongly underresolved and grid-point storms occur (Weisman et al. 1997). Finally, careful inspection of Fig. 5 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 provides a qualitative overview of the precipitation distribution. In order to get a more quantitative view, we compute the location and width of the Pacific and Atlantic ITCZs (see Fig. 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. 2016; Siongco et al. 2014). Location and width are computed by defining precipitation ob-
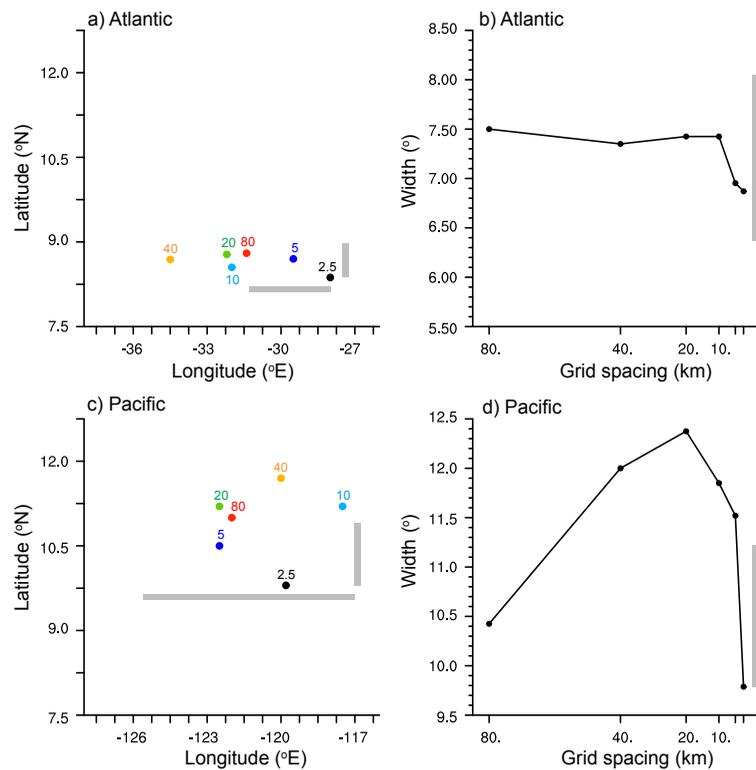
Fig. 6.   (a, c) Location of the ITCZ over the Atlantic and Pacific as well as (b, d) the corresponding width. The gray bar represents the spread of the DYAMOND ensemble (computed as one standard deviation). The GEOS model was not included in the computation of the spread as there were errors in the precipitation diagnostics. Note that all the data have been regridded on a 0.75° grid (approximately the grid spacing of the 80-km simulation) for a fair comparison.

jects as inspired from the SAL (Structure, Amplitude, and Location) measure (Wernli et al. 2008) and, for instance, applied by 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 elliptical objects, give the location and meridional width of the ITCZ. Visual inspection of Fig. 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 spacings tend here to shift the Atlantic ITCZ

eastwards. 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: 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 spacings 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 a higher resolution (Fig. 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 (Figs. 6a, c). However, 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-northeastern flank of the ITCZ (see Figs. 5, 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 the 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-northeast 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., Fig. 4a), the finer-resolution simulations will tend to rain more over low SSTs and less over high SSTs in comparison to the coarser-resolution simulations.

Concerning the ITCZ width (Fig. 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 (Fig. 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, Fig. 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, albeit with a much less pronounced increase in precipitation on the northern flank of the ITCZ when the grid spacing is coarsened. For the sake of clarity, we only show the 2.5-km and 80-km simulations. We also define in Fig. 7 the three cloud categories shallow, congestus, and deep on the basis of a cloud top below 4 km, between 4 km and 8 km, and above 8 km, respectively.

As visible in Fig. 7, both simulations exhibit a placement of the ITCZ in the northern hemisphere with trade winds converging toward the location of
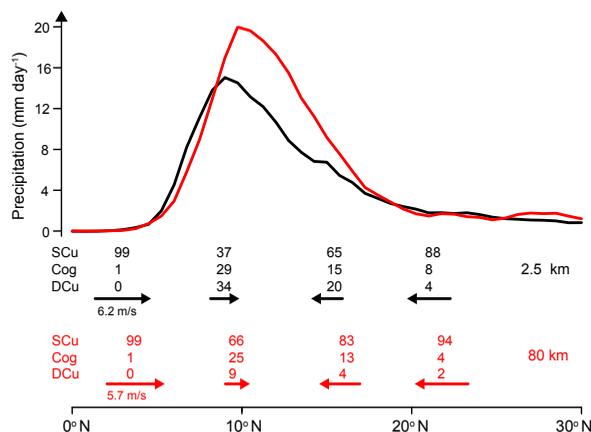
Fig. 7. Picture of the ITCZ with precipitation (mm day^{-1}, line), wind velocity (m s^{-1}, arrow), and cloud distribution (%, number) derived from the results of the 2.5-km (black) and 80-km (red) simulations for the Pacific ITCZ. Only the eastern side of the Pacific, comprised between 170° and 90°W (see the location of the precipitation object in Fig. 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 on the basis of the 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.

the precipitation maximum at the core of the ITCZ. The trade winds reach values between 6 m s^{-1} and 7 m s^{-1} in the subtropics and decelerate to a few meters per second in the doldrums. At the same time, the clouds transition from shallow to deeper clouds, with a successive increase in 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 % at 6° north of it in both simulations. Although a direct comparison to studies that have used satellite observations to quantify cloud partitioning is difficult because of the wide variety of employed cloud definitions, 30 % of all clouds being congestus is larger than what was found by 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 in Fig. 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 at 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 key element of the extratropical circulation is the jet stream. Despite being not directly linked to cloud and convective processes, GCMs using a 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 by our previous analysis on precipitation objects, we diagnose the location of the polar jet in a quantitative manner 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 rather 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. In addition, 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 subtropi-
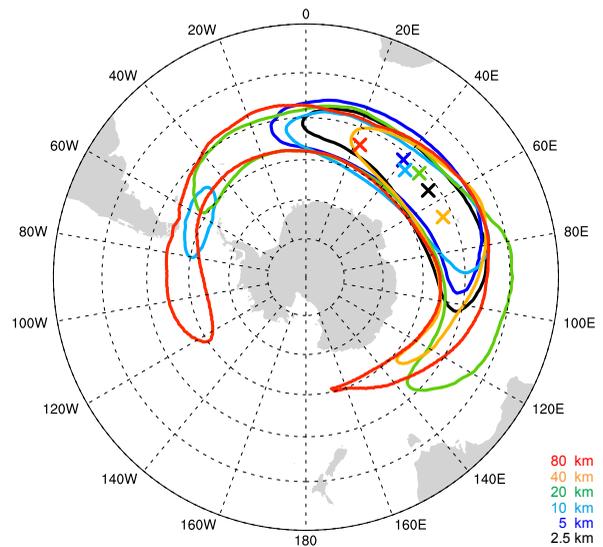
Fig. 8. Location (crosses) and extent (contour lines) of the jet objects for the set of simulations. See the text for more details.

cal jet. The resulting extent of the jet objects for each simulation as well as their centroid is shown in Fig. 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 and 66.4°E. Nevertheless, there seems to be a tendency for more extended jet objects with coarser resolution. 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, with a grid spacing of 1°, the latitudinal position of the polar jet converges toward 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 spacings finer than 50 km.

4. Discussion

The previous section interpreted resolution differences with respect to the spread of the DYAMOND ensemble. From this, we concluded that the resolu-

Table 2.   Various statistics averaged over the simulation period for the 80-km and 2.5-km simulations as well as for the 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). The 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 ITCZ 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° × 0.75° grid spacing. The tropics are from 30°S to 30°N.

Variable	ICON 80 km		ICON 2.5 km		Observation Tropics
	Global	Tropics	Global	Tropics	
Precipitation	3.25	3.90	3.12	3.64	3.37
Latent heat flux	91.8	157.0	89.3	143.5	114.0
Sensible heat flux	28.1	20.4	29.8	17.7	8.2
Surface net longwave radiation	−47.0	−47.1	−56.9	−59.4	−53.7
Surface net shortwave radiation	124.2	149.0	162.4	206.2	205.2
TOA net longwave radiation	−245.3	−265.4	−246.4	−264.3	−260.3
TOA net shortwave radiation	204.2	246.3	240.0	300.0	309.9
Atlantic ITCZ center	8.80, −31.4		8.37, −28.0		8.38, −24.4
Pacific ITCZ center	11.0, −122.0		9.8, −119.8		10.1, −117.2
Atlantic ITCZ width	10.0		9.16		8.21
Pacific ITCZ width	13.9		13.05		12.3
Polar jet center	−53.9, 22.2		−52.2, 54.7		−52.6, 41.4

tion differences between the 5-km and the 2.5-km simulations 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 have not 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 the outgoing longwave radiation on the global scale and over tropical ocean, as well as the 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, for both net shortwave radiation and surface net longwave radiation, signs of convergence are visible with a continuous reduction of the differences between successive grid refinements (see Figs. 1b, c). This is promising as the 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 the observations. This allows us to investigate whether variables exhibiting larger resolution dependencies are associated with larger biases.

The mean values derived from the 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 Fig. 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 a higher resolution brings the simulated values closer to the observed ones. Only for the net shortwave radiation and net longwave radiation at the surface, does refining the grid spacing down to 2.5 km end up overcorrecting the biases originally present at 80 km. The resulting biases in the 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 the 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 the observations.

Our previous analysis also indicated that changes in the grid spacing primarily affect the radiation components, except for the outgoing longwave radiation. The biases for the 2.5-km simulation are -9.9, 1, and -5.7 W m^{-2} for the net shortwave radiation at the top of the atmosphere, at the surface and for the surface net longwave radiation, respectively (Table 2). In comparison, the biases are -4 W m^{-2} for the net longwave radiation at the top of the atmosphere, 9.5 W m^{-2} for the sensible heat flux, 29.5 W m^{-2} for the 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 or dynamics or in the model parameter settings would have a larger impact.

5. Conclusions

In this study, 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, and jet position 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 and the same physical parameterizations, and were not retuned for a specific grid spacing. In order to objectively quantify the resolution dependencies, we compared the resolution-induced differences to the spread obtained in the DYAMOND ensemble of models. The latter comprises eight global climate models integrated at a storm-resolving resolution of O(3 km) for 40 days. Despite the shortness of the simulation period, we believe that the results could hold for a 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 27 statistics and with the exception of the width of the Pacific ITCZ, at least a grid spacing of 5 km appears to be sufficient for capturing 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 many 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, and 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 the 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 the surface net longwave radiation (10 W m^{-2}). These 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 the net shortwave radiation, both at the surface and at the top-of-the-atmosphere, with a finer grid spacing, as well as 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 eastwards with a 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 the simulation biases, but statistics exhibiting less resolution dependency are not necessarily closer to observations.

Even if resolution differences become smaller than the DYAMOND ensemble spread by a grid spacing of 5 km at least, the resolution differences between the 5-km and the 2.5-km simulations are not zero yet, meaning that the investigated climate statistics have not converged in a strict sense. However, the net shortwave radiation, which depicts the largest resolution dependency, shows signs 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 using explicit convection may be used to simulate the climate. Even a grid spacing of 10 km would be sufficient to capture 18 out of the 27 investigated statistics. This is promising given the computational burden associated with simulations using a grid spacing of 2.5 km and seems to confirm the experience with the NICAM model. Only for the width of the Pacific ITCZ and for the fraction of deep convective clouds do large resolution dependencies remain, which could put the use of a grid spacing of 5 km into question.

Acknowledgements

References

Arakelian, A., and F. Codron, 2012: Southern Hemisphere jet variability in the IPSL GCM at varying resolutions. *J. Atmos. Sci.*, **69**, 3788–3799.

Asensio, H., M. Messmer, D. Lüthi, and K. Osterried, 2019: *External parameters for Numerical Weather Prediction and climate application EXTPAR. User and implementation guide*. 44 pp. [Available at http://www.cosmo-model.org/content/support/software/ethz/extpar-userManual-v5.0.pdf.]

Baldauf, M., A. Seifert, J. Förstner, D. Majewski, M. Raschendorfer, and T. Reinhardt, 2011: Operational convective-scale numerical weather prediction with the COSMO model: Description and sensitivities. *Mon. Wea. Rev.*, **139**, 3887–3905.

Ban, N., J. Schmidli, and C. Schär, 2014: Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations. *J. Geophys. Res.*, **119**, 7889–7907.

Bryan, G. H., J. C. Wyngaard, and J. M. Fritsch, 2003: Resolution requirements for the simulation of deep moist convection. *Mon. Wea. Rev.*, **131**, 2394–2416.

Chan, S. C., E. J. Kendon, H. J. Fowler, S. Blenkinsop, C. A. T. Ferro, and D. B. Stephenson, 2013: Does increasing the spatial resolution of a regional climate model improve the simulated daily precipitation? *Climate Dyn.*, **41**, 1475–1495.

Craig, G. C., and A. Dörnbrack, 2008: Entrainment in cumulus clouds: What resolution is cloud-resolving? *J. Atmos. Sci.*, **65**, 3978–3988.

Demory, M.-E., P. L. Vidale, M. J. Roberts, P. Berrisford, J. Strachan, R. Schiemann, and M. S. Mizielinski, 2013: The role of horizontal resolution in simulating drivers of the global hydrological cycle. *Climate Dyn.*, **42**, 2201–2225.

Flato, G., J. Marotzke, B. Abiodun, P. Braconnot, S. Chou, W. Collins, P. Cox, F. Driouech, S. Emori, V. Eyring, C. Forest, P. Gleckler, E. Guilyardi, C. Jakob, V. Kattsov, C. Reaso, and M. Rummukaine, 2013: Evaluation of climate models. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Stocker, T. F., D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley (eds.), Cambridge University Press, Cambridge, UK and New York, USA, 741–866.

Fuhrer, O., T. Chadha, T. Hoefler, G. Kwasniewski, X. Lapillonne, D. Leutwyler, D. Lüthi, C. Osuna, C. Schär, T. C. Schulthess, and H. Vogt, 2018: Near-global climate simulation at 1 km resolution: Establishing a performance baseline on 4888GPUs with COSMO 5.0. *Geosci. Model Dev.*, **11**, 1665–1681.

Giorgetta, M. A., R. Brokopf, T. Crueger, M. Esch, S. Fiedler, J. Helmert, C. Hohenegger, L. Kornblueh, M.

Köhler, E. Manzini, T. Mauritsen, C. Nam, T. Rad-datz, S. Rast, D. Reinert, M. Sakradzija, H. Schmidt, R. Schneck, R. Schnur, L. Silvers, H. Wan, G. Zängl, and B. Stevens, 2018: ICON-A the atmosphere component of the ICON Earth System Model: I. model description. *J. Adv. Model. Earth Syst.*, **10**, 1613–1637.

Hanley, K. E., R. S. Plant, T. H. M. Stein, R. J. Hogan, J. C. Nicol, H. W. Lean, C. Halliwell, and P. A. Clark, 2015: Mixing-length controls on high-resolution simulations of convective storms. *Quart. J. Roy. Meteor. Soc.*, **141**, 272–284.

Heinze, R., A. Dipankar, C. C. Henken, C. Moseley, O. Sourdeval, S. Trömel, X. Xie, P. Adamidis, F. Ament, H. Baars, C. Barthlott, A. Behrendt, U. Blahak, S. Bley, S. Brdar, M. Brueck, S. Crewell, H. Deneke, P. Di Girolamo, R. Evaristo, J. Fischer, C. Frank, P. Frie-derichs, T. Göcke, K. Gorges, L. Hande, M. Hanke, A. Hansen, H.-C. Hege, C. Hoose, T. Jahns, N. Kalthoff, D. Klocke, S. Kneifel, P. Knippertz, A. Kuhn, T. van Laar, A. Macke, V. Maurer, B. Mayer, C. I. Meyer, S. K. Muppa, R. A. J. Neggers, E. Orlandi, F. Pantillon, B. Pospichal, N. Röber, L. Scheck, A. Seifert, P. Seifert, F. Senf, P. Siligam, C. Simmer, S. Steinke, B. Stevens, K. Wapler, M. Weniger, V. Wulfmeyer, G. Zängl, D. Zhang, and J. Quaas, 2017: Large-eddy simulations over Germany using ICON: A comprehensive evaluation. *Quart. J. Roy. Meteor. Soc.*, **143**, 69–100.

Hohenegger, C., P. Brockhaus, and C. Schär, 2008: Towards climate simulations at cloud-resolving scales. *Meteor. Z.*, **17**, 383–394.

Hourdin, F., T. Mauritsen, A. Gettleman, J.-C. Golaz, V. Balaji, Q. Duan, D. Folini, D. Ji, D. Klocke, Y. Qian, F. Rauser, C. Rio, L. Tomassini, M. Watanabe, and D. Williamson, 2017: The art and science of climate model tuning. *Bull. Amer. Meteor. Soc.*, **98**, 589–602.

Ikeda, K., R. Rasmussen, C. Liu, D. Gochis, D. Yates, F. Chen, M. Tewari, M. Barlage, J. Dudhia, K. Miller, K. Arsenault, V. Grubišić, G. Thompson, and E. Guttman, 2010: Simulation of seasonal snowfall over Colorado. *Atmos. Res.*, **97**, 462–477.

Ito, J., S. Hayashi, A. Hashimoto, H. Ohtake, F. Uno, H. Yoshimura, T. Kato, and Y. Yamada, 2017: Stalled improvement in a numerical weather prediction model as horizontal resolution increases to the sub-kilometer scale. *SOLA*, **13**, 151–156.

Jeevanjee, N., 2017: Vertical velocity in the gray zone. *J. Adv. Model. Earth Syst.*, **9**, 2304–2316.

Kajikawa, Y., Y. Miyamoto, R. Yoshida, T. Yamaura, H. Yashiro, and H. Tomita, 2016: Resolution dependence of deep convections in a global simulation from over 10-kilometer to sub-kilometer grid spacing. *Prog. Earth Planet. Sci.*, **3**, 16, doi:10.1186/s40645-016-0094-5.

Klocke, D., M. Brueck, C. Hohenegger, and B. Stevens, 2017: Rediscovery of the doldrums in storm-resolving simulations over the tropical Atlantic. *Nat. Geosci.*, **10**, 891–896.

Langhans, W., J. Schmidli, and C. Schär, 2012: Bulk convergence of cloud-resolving simulations of moist convection over complex terrain. *J. Atmos. Sci.*, **69**, 2207–2228.

Lu, J., G. Chen, L. R. Leung, D. A. Burrows, Q. Yang, K. Sakaguchi, and S. Hagos, 2015: Toward the dynamical convergence on the jet stream in aquaplanet AGCMs. *J. Climate*, **28**, 6763–6782.

Maher, P., G. K. Vallis, S. C. Sherwood, M. J. Webb, and P. G. Sansom, 2018: The impact of parameterized convection on climatological precipitation in atmospheric global climate models. *Geophys. Res. Lett.*, **45**, 3728–3736.

Marsham, J. H., N. S. Dixon, L. Garcia-Carreras, G. M. S. Lister, D. J. Parker, P. Knippertz, and C. E. Birch, 2013: The role of moist convection in the West African monsoon system: Insights from continental-scale convection-permitting simulations. *Geophys. Res. Lett.*, **40**, 1843–1849.

Mass, C. F., D. Ovens, K. Westrick, and B. A. Colle, 2002: Does increasing horizontal resolution produce more skillful forecasts? *Bull. Amer. Meteor. Soc.*, **83**, 407–430.

Mauritsen, T., B. Stevens, E. Roeckner, T. Crueger, M. Esch, M. Giorgetta, H. Haak, J. Jungclaus, D. Klocke, D. Matei, U. Mikolajewicz, D. Notz, R. Pincus, H. Schmidt, and L. Tomassini, 2012: Tuning the climate of a global model. *J. Adv. Model. Earth Syst.*, **4**, M00A01, doi:10.1029/2012MS000154.

Medeiros, B., and B. Stevens, 2011: Revealing differences in GCM representations of low clouds. *Climate Dyn.*, **36**, 385–399.

Miura, H., M. Satoh, H. Tomita, A. T. Noda, T. Nasuno, and S. Iga, 2007: A short-duration global cloud-resolving simulation with a realistic land and sea distribution. *Geophys. Res. Lett.*, **34**, L02804, doi:10.1029/2006GL027448.

Miyamoto, Y., Y. Kajikawa, R. Yoshida, T. Yamaura, H. Yashiro, and H. Tomita, 2013: Deep moist atmospheric convection in a subkilometer global simulation. *Geophys. Res. Lett.*, **40**, 4922–4926.

Miyamoto, Y., R. Yoshida, T. Yamaura, H. Yashiro, H. Tomita, and Y. Kajikawa, 2015: Does convection vary in different cloudy disturbances? *Atmos. Sci. Lett.*, **16**, 305–309.

Mlawer, E. J., and S. A. Clough, 1998: Shortwave and longwave enhancements in the rapid radiative transfer model. *Proceedings of the 7th Atmospheric Radiation Measurement (ARM) Science Team Meeting*, CONF-970365, San Antonio, Texas, 409–413.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102**, 16663–16682.

Neggers, R. A. J., J. D. Neelin, and B. Stevens, 2007: Impact mechanisms of shallow cumulus convection on tropical climate dynamics. *J. Climate*, **20**, 2623–2642.

Noda, A. T., K. Oouchi, M. Satoh, H. Tomita, S. Iga, and Y. Tsushima, 2010: Importance of the subgrid-scale turbulent moist process: Cloud distribution in global cloud-resolving simulations. *Atmos. Res.*, **96**, 208–217.

Noda, A. T., K. Oouchi, M. Satoh, and H. Tomita, 2012: Quantitative assessment of diurnal variation of tropical convection simulated by a global nonhydrostatic model without cumulus parameterization. *J. Climate*, **25**, 5119–5134.

Panosetti, D., L. Schlemmer, and C. Schär, 2018: Convergence behavior of idealized convection-resolving simulations of summertime deep moist convection over land. *Climate Dyn.*, doi:10.1007/s00382-018-4229-9.

Petch, J. C., A. R. Brown, and M. E. B. Gray, 2002: The impact of horizontal resolution on the simulations of convective development over land. *Quart. J. Roy. Meteor. Soc.*, **128**, 2031–2044.

Phillips, T. J., G. L. Potter, D. L. Williamson, R. T. Cederwall, J. S. Boyle, M. Fiorino, J. J. Hnilo, J. G. Olson, S. Xie, and J. J. Yio, 2004: Evaluating parameterizations in general circulation models: Climate simulation meets weather prediction. *Bull. Amer. Meteor. Soc.*, **12**, 1903–1915.

Pope, V. D., and R. A. Stratton, 2002: The processes governing horizontal resolution sensitivity in a climate model. *Climate Dyn.*, **19**, 211–236.

Prein, A. F., A. Gobiet, M. Suklitsch, H. Truhetz, N. K. Awan, K. Keuler, and G. Georgievski, 2013: Added value of convection permitting seasonal simulations. *Climate Dyn.*, **41**, 2655–2677.

Prein, A. F., W. Langhans, G. Fosser, A. Ferrone, N. Ban, K. Goergen, M. Keller, M. Tölle, O. Gutjahr, F. Feser, E. Brisson, S. Kollet, J. Schmidli, N. P. M. van Lipzig, and R. Leung, 2015: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Rev. Geophys.*, **53**, 323–361.

Prein, A. F., A. Gobiet, H. Truhetz, K. Keuler, K. Goergen, C. Teichmamn, C. F. Maule, E. van Meijgaard, M. Déqué, G. Nikulin, R. Vautard, A. Colette, E. Kjellström, and D. Jacob, 2016: Precipitation in the EURO-CORDEX 0.11° and 0.44° simulations: High resolution, high benefits? *Climate Dyn.*, **46**, 383–412.

Qian, J.-H., 2008: Why precipitation is mostly concentrated over islands in the maritime continent. *J. Atmos. Sci.*, **65**, 1428–1441.

Rasmussen, R., C. Liu, K. Ikeda, D. Gochis, D. Yates, F. Chen, M. Tewari, M. Barlage, J. Dudhia, W. Yu, K. Miller, K. Arsenault, V. Grubišić, G. Thompson, and E. Gutmann, 2011: High-resolution coupled climate runoff simulations of seasonal snowfall over Colorado: A process study of current and warmer climate. *J. Climate*, **24**, 3015–3048.

Richard, E., A. Buzzi, and G. Zängl, 2007: Quantitative precipitation forecasting in the Alps: The advances achieved by the Mesoscale Alpine Programme. *Quart. J. Roy. Meteor. Soc.*, **133**, 831–846.

Roberts, N. M., and H.W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97.

Sato, T., T. Yoshikane, M. Satoh, H. Miura, and H. Fujinami, 2008: Resolution dependency of the diurnal cycle of convective clouds over the Tibetan Plateau in a mesoscale model. *J. Meteor. Soc. Japan*, **86A**, 17–31.

Sato, T., H. Miura, M. Satoh, Y. N. Takayabu, and Y. Wang, 2009: Diurnal cycle of precipitation in the tropics simulated in a global cloud-resolving model. *J. Climate*, **22**, 4809–4826.

Satoh, M., T. Matsuno, H. Tomita, H. Miura, T. Nasuno, and S. Iga, 2008: Nonhydrostatic ICosahedral Atmospheric Model (NICAM) for global cloud resolving simulations. *J. Comput. Phys.*, **227**, 3486–3514.

Satoh, M., B. Stevens, F. Judt, M. Khairoutdinov, S.-J. Lin, W. M. Putman, and P. Düben, 2019: Global cloud-resolving models. *Curr. Climate Change Rep.*, **5**, 172–184.

Schiemann, R., M.-E. Demory, M. S. Mizielinski, M. J. Roberts, L. C. Shaffrey, J. Strachan, and P. L. Vidale, 2014: The sensitivity of the tropical circulation and Maritime Continent precipitation to climate model resolution. *Climate Dyn.*, **42**, 2455–2468.

Schrodin, R., and E. Heise, 2002: A new multi-layer soil model. *COSMO Newsl.*, **2**, 149–151.

Schwartz, C. S., J. S. Kain, S. J. Weiss, M. Xue, D. R. Bright, F. Kong, K. W. Thomas, J. J. Levit, and M. C. Coniglio, 2009: Next-day convection-allowing WRF model guidance: A second look at 2-km versus 4-km grid spacing. *Mon. Wea. Rev.*, **137**, 3351–3372.

Sexton, D. M. H., A. V. Karmalkar, J. M. Murphy, K. D. Williams, I. A. Boutle, C. J. Morcrette, A. J. Stirling, and S. B. Vosper, 2019: Finding plausible and diverse variants of a climate model. Part 1: Establishing the relationship between errors at weather and climate time scales. *Climate Dyn.*, **53**, 989–1022.

Siongco, A. C., C. Hohenegger, and B. Stevens, 2014: The Atlantic ITCZ bias in CMIP5 models. *Climate Dyn.*, **45**, 1169–1180.

Stanfield, R. E., J. H. Jiang, X. Dong, B. Xi, H. Su, L. Donner, L. Rotstayn, T. Wu, J. Cole, and E. Shindo, 2016: A quantitative assessment of precipitation associated with the ITCZ in the CMIP5 GCM simulations. *Climate Dyn.*, **47**, 1863–1880.

Stevens, B., M. Satoh, L. Auger, J. Biercamp, C. S. Bretherton, X. Chen, P. Düben, F. Judt, M. Khairoutdinov, D. Klocke, C. Kodama, L. Kornblueh, S.-J. Lin, P. Neumann, W. M. Putman, N. Röber, R. Shibuya, B. Vanniere, P. L. Vidale, N. Wedi, and L. Zhou, 2019: DYAMOND: The DYnamics of the Atmospheric

general circulation Modeled On Non-hydrostatic Domains. *Prog. Earth Planet. Sci.*, **6**, 61, doi:10.1186/s40645-019-0304-z.

Swart, N. C., and J. C. Fyfe, 2012: Observed and simulated changes in the Southern Hemisphere surface westerly wind-stress. *Geophys. Res. Lett.*, **39**, L16711, doi:10.1029/2012GL052810.

Tomita, H., H. Miura, S. Iga, T. Nasuno, and M. Satoh, 2005: A global cloud-resolving simulation: Preliminary results from an aqua planet experiment. *Geophys. Res. Lett.*, **32**, L08805, doi:10.1029/2005GL022459.

Vannière, B., M.-E. Demory, P. L. Vidale, R. Schiemann, M. J. Roberts, C. D. Roberts, M. Matsueda, L. Terray, T. Koenigk, and R. Senan, 2019: Multi-model evaluation of the sensitivity of the global energy budget and hydrological cycle to resolution. *Climate Dyn.*, **52**, 6817–6846.

Wall, C., C. Liu, and E. Zipser, 2013: A climatology of tropical congestus using CloudSat. *J. Geophys. Res.*, **118**, 6478–6492.

Webb, M. J., A. P. Lock, C. S. Bretherton, S. Bony, J. N. S. Cole, A. Idelkadi, S. M. Kang, T. Koshiro, H. Kawai, T. Ogura, R. Roehrig, Y. Shin, T. Mauritsen, S. C. Sherwood, J. Vial, M. Watanabe, M. D. Woelfle, and M.

Zhao, 2015: The impact of parameterized convection on cloud feedback. *Philos. Trans. Roy. Soc. A*, **373**, 20140414, doi:10.1098/rsta.2014.0414.

Weisman, M. L., W. C. Skamarock, and J. B. Klemp, 1997: The resolution dependence of explicitly modeled convective systems. *Mon. Wea. Rev.*, **125**, 527–548.

Weisman, M. L., C. Davis, W. Wang, K. W. Manning, and J. B. Klemp, 2008: Experiences with 0-36-h explicit convective forecasts with the WRF-ARW model. *Wea. Forecasting*, **23**, 407–437.

Wernli, H., M. Paulat, M. Hagen, and C. Frei, 2008: SAL—A novel quality measure for the verification of quantitative precipitation forecasts. *Mon. Wea. Rev.*, **136**, 4470–4487.

Yashiro, H., Y. Kajikawa, Y. Miyamoto, T. Yamaura, R. Yoshida, and H. Tomita, 2016: Resolution dependency of the diurnal cycle of precipitation simulated by a global cloud-system resolving model. *SOLA*, **12**, 272–276.

Zängl, G., D. Reínert, P. Ripodas, and M. Baldauf, 2015: The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core. *Quart. J. Roy. Meteor. Soc.*, **141**, 563–579.