Assessing precipitation over the Amazon basin as simulated by a storm-resolving model

L. Paccini[1], B. Stevens[1]

[1]Max Planck Institute for Meteorology
[1]Hamburg, Germany

Key Points:

- The explicit representation of convection and organized convective systems (OCS) enable improvements in the simulation of Amazon rainfall.
- Surface processes influence the propagation of diurnal OCS and strong low-level easterlies are related to the occurrence of nocturnal OCS.
- Outstanding biases show insensitivity to two fold refinement in horizontal mesh, indicative of the importance of much smaller scale processes.

Corresponding author: Laura Paccini, `laura.paccini@mpimet.mpg.de`

Abstract

In this study we investigate whether a better representation of precipitation in the Amazon basin arises through an explicit representation of convection and whether it is related to the representation of organized systems. In addition to satellite data, we use ensemble simulations of the ICON-NWP model at storm-resolving (2.5 km to 5.0 km) scales with explicit convection (E-CON) and coarse resolutions, with parameterized convection (P-CON). The main improvements in the representation of Amazon precipitation by E-CON are in the spatial pattern of precipitation, the distribution of precipitation intensity and the spatial distribution in the diurnal cycle. By isolating precipitation from organized convective systems (OCS), it is shown that many of the well simulated precipitation features in the Amazon arise from the distribution of these systems. The simulated and observed OCS are classified into 6 clusters which distinguish nocturnal and diurnal OCS. While the E-CON ensembles capture the OCS, especially their diurnal cycle, their frequency is reduced compared to observations. Diurnal clusters are influenced by surface processes such as cold pools, which aid to the propagation of OCS. Nocturnal clusters are rather associated with strong low-level easterlies, possibly related to the Amazonian low-level jet. These particular environmental conditions provide insights on the processes that are important for OCS in the Amazon and should be further improved.

Plain Language Summary

The Amazon basin is a relevant element of the Earth system since it influences the global water and carbon cycle, as well as it constitutes a unique ecosystem. Over this important region, conventional climate models do not simulate basic features of rainfall given their inability to resolve this physical process due to their coarse spatial resolution. In this study, we use high-resolution simulations that allow an explicit representation of such physical process (moist convection) and compare them with a set of coarse-resolution simulations and observations. We find that improvements in the representation of Amazon rainfall, such as the distribution of light and high intensity rain rates, as well as the spatial variability of the diurnal cycle, are explained by the explicit representation of moist convection. Moreover, these improvements arise from the representation of big and organized systems that produce intense rainfall (OCS). We find that particular environmental conditions are associated with the OCS according to their time of occurrence. Diurnal OCS are mainly influenced by interactions with the surface, while nocturnal OCS are related to strong low-level winds.

1 Introduction

The Amazon basin is the largest rainforest in the Earth and of great relevance for the global hydro-climate and biodiversity (Marengo, 2006; Phillips et al., 2008). It is also a region, like many in the tropics, where climate model precipitation biases are both large and systematic. These biases are evident in every aspect of the representation of precipitation, from its spatial and temporal distribution, to its intensity and form. Models systematically have too little precipitation over the northern Amazon (e.g., Yin et al., 2013; Fiedler et al., 2020). The diurnal cycle is characterized by a too early precipitation peak (Betts & Jakob, 2002; Tang et al., 2021) and evidence of convective organization (Mapes & Neale, 2011), which has been estimated to account for up to 50 % of the total Amazon rain (Feng et al., 2021), is effectively absent. In this study, we use kilometer-scale "storm-resolving" simulations over large domains to assess the degree to which they reduce these biases and the extent to which this depends on the explicit representation of organized convective systems. In doing so our premise is that convective features which are not improved, or for which remaining biases show no clear sign of improvement with increases in resolution, are indicative of an important role for non-convective, e.g., cloud microphysical, small (sub hectometer) scale mixing, or land-surface processes.

The distinguishing characteristic of storm-resolving models is that they explicitly represent the transient dynamics of convective storm systems, whose length-scales are commensurate with the depth of the troposphere (Satoh et al., 2019). Representing these features become possible at grid spacings of 5 km to 10 km although there is considerable evidence that convection is increasingly distorted as grid spacings increase above 1 km to 2 km. Nonetheless, the ability to represent convective entities as geometric objects that interact dynamically with their environment, and are governed by the correct physical relations (laws of motion) seems to explain why even on 5 km to 10 km gird meshes, an explicit representation of convection leads to more physical representations of convection than what is possible using convective parameterization (e.g., Love et al., 2011; Birch et al., 2015). In recent years, storm-resolving models have shown systematic improvements in representing precipitation, albeit to a degree that seems to vary from place to place. For instance, Arnold et al. (2020) found regional differences in the mean precipitation of a 40-day global simulation, where precipitation is overestimated over Africa but underestimated over the Great Plains in North America. While the shortness of the simulations (40 days) and the remote influence of larger-scale biases might explain this discrepancy, they also found that precipitation tends to peak earlier than observations over regions dominated by local thermodynamic forcing; whereas the largest improvements were found in regions where the diurnal cycle is driven by non-local propagating convection.

Although storm-resolving models overcome the long-standing "drizzle" problem of convective parameterizations (Stephens et al., 2010), they still disagree in the representation of high intensity precipitation rates (>80 mm d^{-1}), which is strongly overestimated (Becker et al., 2021) in some models, and apparently underestimated in others (Arnold et al., 2020; Judt & Rios-Berrios, 2021). High intensity precipitation can be related to organized convective systems which we expect storm-resolving models to better represent as compared to models dependent on parameterized convection (e.g., Stevens et al., 2020). How much improvements in the representation of precipitation relate to the representation organized convection is not evident and has not been investigated yet.

A few studies have begun evaluating the representation of precipitation over the Amazon basin using global storm-resolving models. For example, (Inoue et al., 2021) compared the semi-diurnal cycle of precipitation with observations for a 5-day period at 3.5 km grid spacing. They found that the model captures the semi-diurnal variation of precipitation in the Amazon basin but it tends to overestimate their amplitudes, especially the second peak during the early morning. Arnold et al. (2020) also analyzed a set of global simulations and found a larger simulated amplitude than observed at a reduced grid spacing (3.5 km). However, in contrast to Inoue et al. (2021), their model did not capture the phase of the precipitation diurnal cycle in the Amazon.

For the most part, storm-resolving model simulations at the regional scale have not been able to look at precipitation over the Amazon in its entirety. For instance, Santos et al. (2019) used a small domain enclosing the city of Manaus and a grid scale of about 780 m. They found that seasonal floods can enhance the intensity of river circulations during daytime and hence convection. Over the eastern Amazon at a grid spacing of 1.5 km, Herbert et al. (2021) analyzed the impact of biomass burning on the diurnal cycle of precipitation. They found that convection is suppressed in the afternoon but enhanced overnight due to aerosol-radiation interactions. Another recent study by Tai et al. (2021), investigated the influence of data assimilation on regional modeling of Amazon precipitation. They performed a 30-day simulation at 4 km grid spacing and focused the analysis on the central Amazon. Their study highlights the improved representation of spatial variability in the precipitation diurnal cycle in contrast to a standard climate model. This feature is related to the representation of organized convective systems which are absent in models reliant on convective parameterization.

In this study we perform storm-resolving simulations with the ICON model over a large domain to study the representation of precipitation over the Amazon. In contrast to previous regional modeling studies, we focus on the Amazon basin in its entirety and, unlike month-long simulations with global models, we perform an ensemble of 30-day simulations at 2.5 km and 5 km grid-spacing. Simulations are performed during March as this is the month with the largest convective activity (e.g., Rehbein et al., 2018). For this period we document the ability of storm-resolving simulations to capture the multi-faceted properties of precipitation as observed over the Amazon, in comparison with a model that arguably uses the most efficient, and certainly well calibrated, statistical representation of convection, i.e., that developed by Bechtold (2017) for the Integrated Forecast System of the European Centre for Medium-range Weather Forecasts. We especially focus on the role of organized convection in improving the representation of precipitation and the extent to which this is coupled to particular environmental conditions. By using two resolutions we further infer to what extent remaining deficits in the representation of precipitation are likely to be improved by modest (factor of 2) refinements in resolution. This question becomes interesting in light of proposals to develop climate information systems based on global models with grid meshes of roughly 1 km (Slingo et al., 2022) as it helps identify the problems that km-scale global models are likely to solve, and those whose solution might require improvements in the representation of processes that remain unresolved, or severely distorted, even on global km-scale meshes.

2 Data and methodology

2.1 Observations

We use the Climate Prediction Center Morphing Method (CMORPH; Xie et al., 2017) dataset for the period from 2010 to 2019. This product estimates precipitation based on passive microwave instruments. The main advantages of CMORPH data are its high temporal (30min) and spatial (8 km) resolutions. Previous studies have also validated its good performance over the Amazon region (e.g., Janowiak et al., 2005). We also compared the analysis with other high-resolution datasets and similar results were obtained; therefore we chose the CMORPH data.

2.2 CMIP6

We use simulations from the Coupled Model Inter-comparison Project: Phase 6 (CMIP6; Eyring et al., 2016). Multi-model ensemble means are used from the historical simulations of the 21th century (2000-2014) and are the same used in Fiedler et al. (2020). We use daily and 3-hourly data available from 14 and 13 models, respectively. Simulations were spatially interpolated to the common T63 grid (about 180 km), the native grid of MPI-ESM low-resolution configuration. For a detailed list of the models, the reader is referred to the supplementary material of Fiedler et al. (2020).

2.3 ICON-NWP

We use the Icosaedral Nonhydrostatic (ICON) atmospheric model (Zängl et al., 2015) in the numerical weather prediction (NWP) configuration. Among the applied physical parameterizations by this model, the parameterization of moist convection is only used for the coarser grid spacing in our experiments. It consists of a bulk mass-flux scheme (Bechtold, 2017), which is one of the latest implementations in the NWP of European meteorological services. Parameterizations common to all simulations are given for processes such as radiation, microphysics and turbulence as described in Zängl et al. (2015). Also, the ICON-NWP model uses the multi-layer land-surface scheme TERRA (Heise et al., 2006).

As initial conditions for the simulations we use the operational analysis data from the European Centre for Medium-Range Weather Forecasts (ECMWF) - Integrated Forecast System (IFS), and from the Hadley Centre Sea Ice and Sea Surface Temperature Center (HadISST; Rayner et al., 2003) for SST. Grids and external parameters (e.g. land properties, topography) are retrieved from the Online Grid Generator tool from the German Meteorological Service (DWD).

2.3.1 Experimental set-up

We conduct a set of simulations using the same approach as Paccini et al. (2021). Global simulations, at 40 km grid spacing (P-CON simulations), serve as initial and boundary conditions to the one-way nested domains at finer grid spacing. The three inner domains have the convective parameterization switched off (E-CON simulations) and comprise the same regions as described in Paccini et al. (2021). The horizontal resolution is successively increased from 20 km to 10 km and to 5 km, with the finest grid spacing covering the tropical Atlantic sector (85°W-25°E; 25°S-25°N). In all domains the vertical resolution includes 90 levels, with the model top at 75 km.

We start 8 simulations at the beginning of March, with different atmospheric states but with the same fixed sea surface temperature (SST), which does not vary over time. Simulations are integrated for 40 days and the analysis is performed over the last 31 days, representing the simulation of March.

We conduct another set of simulations using an updated version of ICON (v2.6.01) with an additional inner domain, at a grid spacing of 2.5 km, that bounds the region: 81°W-36°W; 21°S-11°N. Given the high computational demands, only 2-member simulations are performed.

In our analysis we compare the 8-member ensemble of P-CON and E-CON at 40 km and 5 km, respectively, with the 2-member ensemble of E-CON at 2.5 km. Although from different ensembles, the E-CON simulations at 2.5 km and 5 km lead to the same results as those E-CON at 2.5 km and 5 km from the 2-member ensemble. We present then results of the 8-member E-CON simulations at 5 km due to more robust statistics.

All data and simulation outputs are regrided to the resolution of the P-CON experiments (about 40 km) except for the CMIP6 ensemble which keeps the grid spacing of about 180 km. The CMIP6 data only serves as a reference of how state-of-the-art climate models, representing the average convective parameterizations, simulate Amazon precipitation.

3 Representation of precipitation

3.1 Geographic distribution

One of the basic metrics when evaluating the representation of rainfall is the mean amount of precipitation, and its spatial pattern. The prevailing bias in most climate models is the underestimation of rain in the Amazon, especially during the wet season (Fiedler et al., 2020). Spatially, the bias shows up as enhanced rain over the eastern region of Brazil and insufficient rain in the central Amazon (Fig. 1, e). This is a bias that does not appear to be related to a poor representation of SST patterns in coupled models, as it has also been documented in simulations using prescribed SST (Richter & Xie, 2008).

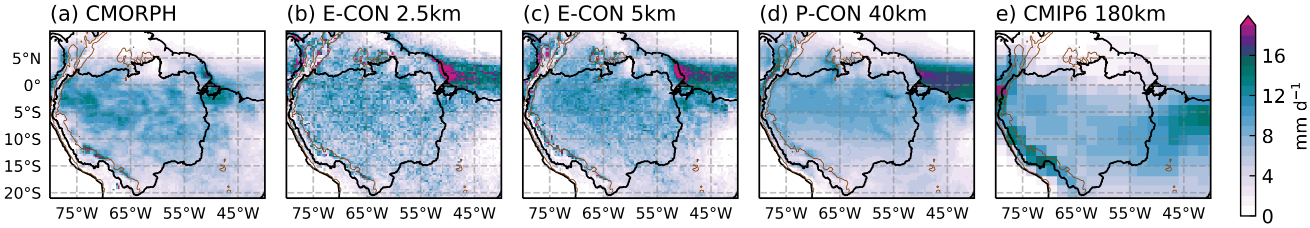

Figure 1. Mean precipitation in March from (a) CMORPH observations, simulations with explicit (E-CON) convection at (b) 2.5 km, (c) 5 km, parameterized convection (P-CON) at (d) 40km and the (e) CMIP6 multi-model ensemble mean. Data and output simulations are regrided to 40km except for CMIP6 models which were interpolated to a common grid of about 180km and only serves as a reference. The Amazon basin is defined as black contours and the topography at 1000m, in brown contours.

Both E-CON and P-CON display a better representation of the mean spatial pattern compared to the CMIP6 ensemble, meaning more rain over the central Amazon and less over eastern Brazil (Fig. 1, b, c, d). However, it still appears that the simulated precipitation is underestimated compared to the observed climatology. This can be partly explained by the simulated broader precipitation band with enhanced rainfall north of the Amazon in both E-CON and P-CON, probably related to the invariable SST used in the simulations. As a result, less precipitation in E-CON and P-CON than CMORPH is observed south of 5°S. In the case of the P-CON ensemble, the spatial distribution is more uniform, with no regions having precipitation rates larger than 12 mm d^{-1}. The E-CON ensembles do show sub-regions with larger mean values, similar to CMORPH, but rainfall over parts of the western Amazon is still underestimated. This characteristic appears insensitive to modest changes in the grid spacing, as evidenced by the similarity between the 2.5 km and 5 km E-CON ensembles (Fig. 1, b, c).

A more localized feature, which appears sensitive to the treatment of convection, is the coastal precipitation over the northeastern coast of Brazil. Simulations with parameterized convection (P-CON and CMIP6) show a lack of precipitation in this region, a bias that is not evident in the E-CON ensembles. Having an adequate representation of the coastal precipitation is thought to be important for the Amazon, due to organized convective systems that originate there and propagate inland (e.g., Greco et al., 1990). Improvements in the representation of coastal precipitation with explicit convection might be related to a better representation of breeze circulations and/or the transition from shallow to deep convection.

A quantitative comparison is presented in Tab. 1. Precipitation is averaged over the Amazon basin and the continental region comprised by 20°S-10°N; 80°W-38°W. Even though the differences among simulations are relatively small ($<1\,\mathrm{mm\,d^{-1}}$), this analysis suggests that the E-CON simulations better match the observations, increasingly so with finer grid spacing, and in regions of less orographic relief (regions below 1000 m above sea-level). The E-CON ensembles differ from observed values by less than $0.35\,\mathrm{mm\,d^{-1}}$ while precipitation biases of the P-CON simulations are nearly twice as large ($>0.5\,\mathrm{mm\,d^{-1}}$).

Table 1. Averaged precipitation over the Amazon Basin (AB) and the ratio of Amazon and tropical South America (SA, 20°S-10°N; 80°W-38°W) rain rates. Values in parentheses are the averages over regions were topography is below 1000 m. For these calculations, observations and output simulations were spatially interpolated onto the CMIP6 grid (180 km).

Dataset	Mean precipitation AB ($\mathrm{mm\,d^{-1}}$)	Ratio AB/SA
CMORPH	7.86 (8.07)	1.28
E-CON 2.5km	7.71 (7.88)	1.21
E-CON 5km	8.19 (8.42)	1.19
P-CON 40km	7.33 (7.52)	1.17
CMIP6 180km	7.82 (7.41)	1.08

Comparing the ratio between Amazon precipitation and the tropical continent as a whole, both E-CON and P-CON display a ratio of about 1.2, similar to observations, whereas the CMIP6 ensemble shows a value closer to 1. This is related to the enhanced precipitation over high topography and over the eastern coast of Brazil in the CMIP6 ensemble. Although the improvements are small, in all measures the most highly resolved E-CON simulations are closest to the observations.

Some aspects of the simulations show less indication of improving with a reduction of the grid spacing at storm-resolving scales. Whilst all simulations with explicit convection do a fair representation, one might expect a better performance for the 2.5 km mesh simulations as compared to those with a 5.0 km mesh. For instance, along the eastern flank of the Andes (from 10°S - 17°S), the Amazon comprises some of the rainiest places in the region, exhibiting features known as "precipitation hot spots" (e.g., Chavez & Takahashi, 2017). The E-CON ensembles exhibit a similar zonal gradient that maximizes eastward; however not as prominently as is seen in observations. The origin of such precipitation hot-spots is not very clear, but it was found that they comprise convective and stratiform rain (Chavez & Takahashi, 2017). The results suggest that the representation of these precipitation maxima may depend on yet smaller scale orographic features, as a microphysical origin of such localized features is difficult to rationalize. .

3.2 Frequency and intensity

The E-CON ensembles show a notable improvement in the estimated frequency and distribution of precipitation intensity in the Amazon basin (Fig. 2). The frequency of daily precipitation follows the spatial pattern of the mean precipitation (Fig. 1), featuring regions where it rains up to 80 % of the days in E-CON and observations (Fig. 2, a, b and c). The E-CON ensembles also distinguish more rain frequency over land areas than rivers, such as the Amazon river mouth and the Tapajos river (Fig.2 b, c), although details are smoothed by the interpolation to the common analysis grid.

A very different picture is displayed by simulations with parameterized convection, as P-CON shares the biases of the CMIP models, which tend to overestimate the frequency of light rain (Stephens et al., 2010) regardless of the spatial resolution (Fig. 2 d,e). Regions where the mean precipitation is greater or equal than $5\,\mathrm{mm\,d^{-1}}$ (Fig. 1)

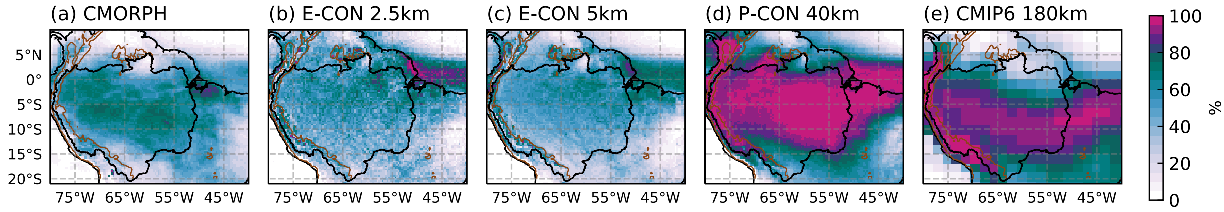

Figure 2. Frequency (%) of daily precipitation greater than $1\,\mathrm{mm\,d^{-1}}$ in March from (a) CMORPH observations and simulations with explicit (E-CON) convection at (b) 2.5 km, (c) 5 km, parameterized convection (P-CON) at (d) 40 km and the (e) CMIP6 multi-model ensemble mean. Data is regrided to 40 km except for CMIP6 models which were interpolated to a common grid of about 180km and only serves as a reference. The Amazon basin is defined as black contours and the topography at 1000m, in brown contours.

show a frequency greater than 90 % to 95 %, indicating that the mean precipitation amount is related to the persistence of rainy days.

To have a broader view of the frequency spectra, Figure 3 displays the distribution of precipitation intensity over the Amazon basin. The E-CON ensembles show an important improvement in the representation of this precipitation feature as compared to simulations with parameterized convection and in agreement with studies focused on different regions (e.g., Holloway et al., 2012; Becker et al., 2021; Judt & Rios-Berrios, 2021). This improvement is evident across the E-CON ensembles, which suggests that it is determined by the treatment of convection rather than the details of the spatial resolution and the experimental set-up (global versus nested, not shown). In a recent comparison study, Judt and Rios-Berrios (2021) showed that simulations with full convective parameterization run at about 4 km grid spacing displayed the same distribution of precipitation intensity as those at 100 km.

Differences in the intensity spectrum are most evident in two intensity intervals. First, the interval between $2\,\mathrm{mm\,d^{-1}}$ to $20\,\mathrm{mm\,d^{-1}}$ (light-to-moderate rain) occurs more frequently, with a clearly preferred intensity in simulations with parameterized convection. Observations and the E-CON ensembles show a flatter distribution, and less frequent rainfall in this intensity interval as a whole. The second intensity interval covers precipitation rates greater than $25\,\mathrm{mm\,d^{-1}}$ (high intensity rain). As compared to observations and to E-CON, these high-intensity rain events are much rarer in P-CON. The

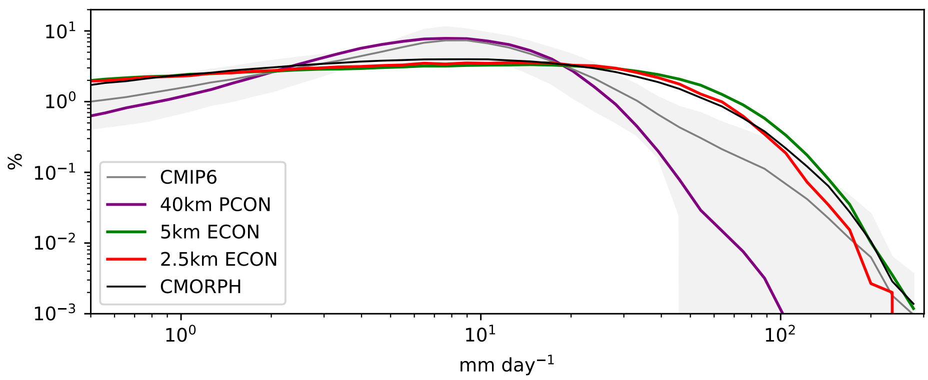

Figure 3. Distribution (%) of daily precipitation intensity greater than 0 mm over the Amazon basin for observations (black line) and simulations (colored lines). Values are binned in a logarithmic scale. The gray shading represent the standard deviation of 14-models of the CMIP6 ensemble.

inter-model variability in the representation of intense precipitation is large across the CMIP models, showing a larger frequency of the multi-model ensemble mean than the P-CON ensemble but still well bellow what is observed. The persistence of this too frequent and too gentle bias (Stephens et al., 2010; Fiedler et al., 2020; Judt & Rios-Berrios, 2021) in all simulations employing parameterized convection suggests that it is not easily addressed in the framework of existing convective parameterizations. The considerably better agreement between observations and simulations that represent convection explicitly, suggests that linking precipitation development to convective motion fields places physical and meaningful constraints on the intensity distribution in ways that parameterizations of convection are unable to mimic.

3.3 Diurnal cycle

The diurnal cycle of precipitation over the Amazon is not spatially homogeneous. To illustrate this feature, we compute the hourly mean for each grid point and then select the time when precipitation is maximum (Fig. 5). This method allows us to consider semidiurnal variations and avoid ambiguities that arise when using the first harmonic approach (Yang et al., 2008).

A large part of the Amazon basin depicts a precipitation maxima in the afternoon from 15 to 18 Local Time (LT) as a result from daytime heating. The afternoon peak

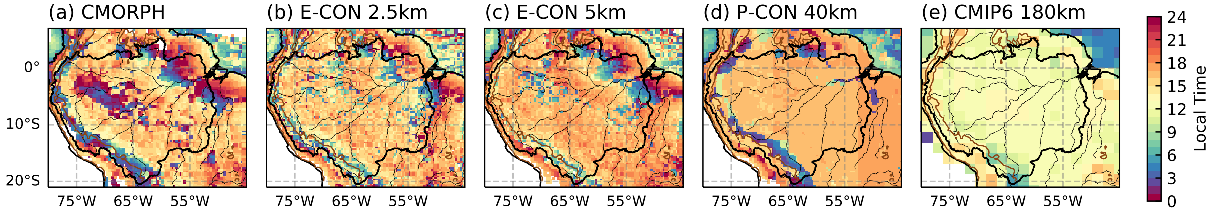

Figure 4. Local time (hour) of maximum precipitation in (a) CMORPH observations and simulations with explicit (E-CON) convection at (b) 2.5 km, (c) 5 km, parameterized convection (P-CON) at (d) 40km and the (e) CMIP6 multi-model ensemble mean. Observations and model outputs are regrided to 40km except for CMIP6 models which were interpolated to a common grid of about 180km and only serves as a reference. In all cases, the hourly mean was calculated and smoothed using a second order Fourier transform per grid point. The Amazon basin is defined as black contours as well as the rivers, and the topography at 1000 m is shown in brown contours.

is reasonably well represented by E-CON and P-CON, although in the case of the latter elements of the parameterization were specifically designed to capture this effect (Bechtold et al., 2008). Nonetheless, it shows that such delays can be represented in the framework of convective parameterization, and thus constitutes an important improvement, but one that apparently has yet to find its way to the CMIP6 multi-model ensemble (Fig.4, e), as these models still tend to precipitate too early (Fiedler et al., 2020; Tang et al., 2021). Perhaps due to the way in which it was implemented, the P-CON simulations displays a rather homogeneous spatial distribution of the time of maximum precipitation. The E-CON simulations, on the contrary, are able to reproduce observed spatial heterogeneities in the diurnal cycle naturally. The time of diurnal precipitation maxima varies between 15 to 18 LT, albeit the 5 km E-CON ensemble displays predominantly a peak time closer to 18LT in contrast to the 2.5 km ensemble.

The spatial heterogeneity of the diurnal cycle in the E-CON ensemble shows a structure that is also evident in the observations. Notable in this respect is the consecutive peaking times from the northeast coast moving inland towards the Amazon. Near the coast, precipitation maximizes close to midday (12-14LT), a feature that may be related to relatively shallow and unorganized convection (e.g., Houze Jr et al., 2015). Precipitation maximizing later in the day, increasingly so as one moves inland, is in agreement with what would be expected from transition to deep convection that propagates towards the Amazon (Greco et al., 1990; Burleyson et al., 2016). The representation of such progressive peaking times and corresponding increasing cloud depth (not shown) suggest that the E-CON ensembles are able to reproduce a realistic transition of convection.

Notwithstanding the general tendency of precipitation to maximize during the day, there are places where precipitation peaks overnight (Garreaud & Wallace, 1997; Rickenbach, 2004; Janowiak et al., 2005; Tanaka et al., 2014). Two regions stand out in CMORPH data displaying a horse-shoe pattern (Fig. 4, a). This structure is captured by the E-CON ensembles (Fig.4, b and c) but is not observed in P-CON or CMIP6. For instance, the northeast extreme of the Amazon basin exhibits a coast-parallel band of consecutive peaking times from 21LT to 6LT (Fig.4, a, b and c). This nocturnal precipitation band has been associated with squall lines, which can originate at the coast and move inland (e.g., Garstang et al., 1994). Other places displaying nocturnal precipitation peaks are not as pronounced in the E-CON ensembles as in observations, but still can be distinguished inland between 6°W-75°W, 5°S-0°W and over the southeast Amazon (50°W-55°W, 15°S-10°S). Many of the nocturnal precipitation peaks also co-locates with the Amazon river and its tributaries, suggesting a sensitivity to the representation of thermally-driven local circulations (e.g., Fitzjarrald et al., 2008; Tanaka et al., 2014; Wu et al., 2021). Over the eastern flank of the Andes, particularly south of 10°S, a nocturnal peak in precipitation is apparently captured by the P-CON ensemble. Closer inspection shows an eastward misplacement of these systems in P-CON, whereas they are better captured by the E-CON ensembles, increasingly so as the grid is refined.

In terms of the amplitude of the precipitation diurnal cycle, Figure 5 shows that both E-CON ensembles particularly overestimate precipitation associated with deep convection (at about 15-17LT). However, in contrast to the results of Inoue et al. (2021), the secondary peak in the early morning is slightly underestimated and rather delayed by 3 hours in both E-CON ensembles. There is not considerable differences between the 2.5 km and 5 km ensembles regarding the amplitude, but only in the phase as shown in Fig. 4.

The representation of the diurnal cycle in the Amazon basin thus proves to be another major area susceptible to the more physical constraints associated with an explicit representation of convection, a finding that is in agreement with a recent study by Tai

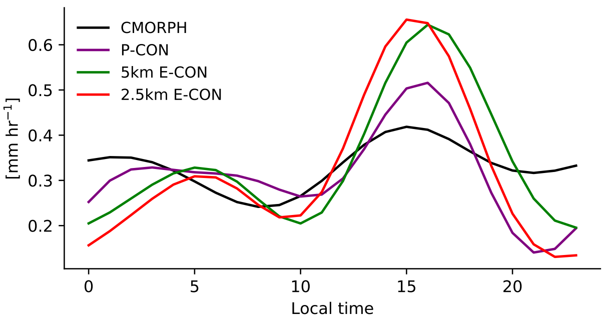

Figure 5. Diurnal precipitation averaged over the Amazon basin.

et al. (2021). Moreover, differences between the 2.5 km and 5 km are more prominent in the time of maximum precipitation, which suggests an improvement with a reduced grid spacing for a daytime precipitation maxima and nocturnal precipitation maxima along the Andes.

4 Role of organized convective systems

In section 3 we compared some precipitation characteristics between observations and small ensembles of simulations, differing in their treatment of convection and in their spatial resolution. An explicit representation of convective precipitation is shown to improve the representation of Amazon precipitation, most notably in terms of the distribution of precipitation intensity and the spatial heterogeneity of the diurnal cycle. These precipitation characteristics can be related to organized convective systems, which develop during the day and can last overnight generating very intense rainfall episodes (e.g., Garreaud & Wallace, 1997; Rickenbach, 2004; Pereira Filho et al., 2015; Rehbein et al., 2018).

In this section we analyze whether improvements in the representation of the precipitation intensity and diurnal cycle by the E-CON ensembles are related to the representation of organized convective systems in the Amazon. Since simulations with parameterized convection fail in reproducing such precipitation features (i.e. high intensity rain rates and nocturnal precipitation peaks), we exclude them from further analysis.

In the following subsections we examine precipitation characteristics of precipitation objects and compare them with the non-organized precipitation. To define a precipitation object, or what we call an organized convective systems (OCS) we use an object-based approach. First, precipitation is associated with grid cells whose hourly rain rate is equal to or greater than $2\,\mathrm{mm\,h^{-1}}$. Precipitation objects are then identified as contiguous grid cells (8-way connection) with a minimum size of $10\,000\,\mathrm{km^2}$ (equivalent to six grid cells on the coarsened analysis grid) at each hour. Given that we do not track the OCS, this method preferentially samples mature systems. This is why we chose a size threshold similar to the mean size found in past studies (Rehbein et al., 2018; Anselmo et al., 2021), which are about $14\,000\,\mathrm{km^2}$ (based on brightness temperature). Even so, we test our findings by redoing the analysis using different thresholds and this did not change our findings.

4.1 Frequency of intensity and size

Figure 6 (a) shows the distribution of precipitation intensity of OCS only (solid lines) and non-organized precipitation (dashed lines). By comparing Fig. 6 (a) with Fig. 3 one can notice a better agreement between the E-CON ensembles and CMORPH data when only considering the OCS. In particular, the 5 km ensemble fits well the observations between $10\,\mathrm{mm\,d^{-1}}$ to $200\,\mathrm{mm\,d^{-1}}$.

Precipitation associated with OCS explains the high-intensity rates ($>100\,\mathrm{mm\,d^{-1}}$) in observations. The distribution of non-organized precipitation in CMORPH resembles the P-CON distribution in Fig. 3. In contrast, non-organized precipitation in the E-CON ensembles still shows larger frequencies of intense rain (around $200\,\mathrm{mm\,d^{-1}}$). This shows a tendency of the E-CON simulations to produce more intense isolated events, which is an expected deficiency at kilometer-scale resolutions given that convection is not fully resolved (Prein et al., 2015; Arnold et al., 2020).

The relative contribution of OCS to the total rainfall can be associated with the distributions of precipitation intensity. In observations, most of the intense precipitation ($>50\,\mathrm{mm\,d^{-1}}$) is associated with OCS (e.g., Feng et al., 2021). The contribution of OCS in the E-CON ensembles is not as large as observed (30 % in the simulations as compared to about 50 % in the observations, not shown), a bias that may arise because high

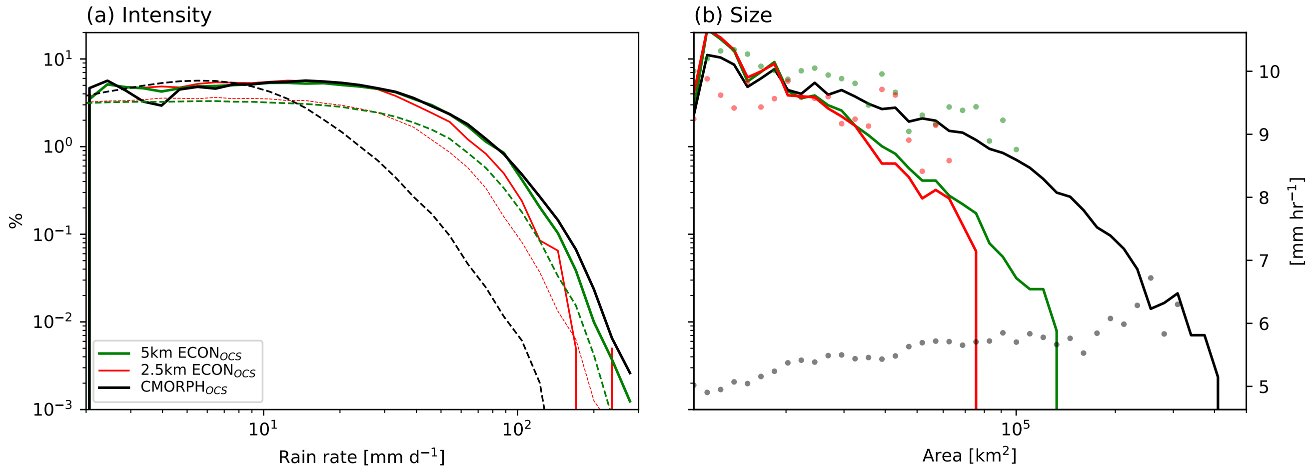

Figure 6. (a) Distribution (%) of daily precipitation intensity of organized convective systems (OCS, solid lines) and non-organized precipitation (dashed lines) in the Amazon basin. (b) Size distribution of organized convective systems (solid lines) and mean precipitation per area bin (scatter points, right axis). Observations are displayed in black and the E-CON ensembles, in green (5 km) and red (2.5 km) colors. Values of intensity and size are binned in a logarithmic scale.

intensity rates are present in the non-organized precipitation events to a greater degree than in the observations.

Another feature related to the precipitation intensity in OCS is their size (Fig. 6, b). As found in some previous studies (e.g., Crook et al., 2019; Arnold et al., 2020), the storm-resolving simulations generally produce smaller precipitation clusters than those identified in observations. The median size for 5 km and 2.5 km E-CON ensembles are 14 411 km^2 and 14 371 km^2, respectively; whereas for CMORPH it is 19 224 km^2. Likewise, the median intensity per bin size is about twice in E-CON than CMORPH (colored dots in Fig. 6, b). The size distribution of OCS shows that E-CON overestimates the frequency of systems smaller than <20 000 km^2 and misses those larger than 150 000 km^2.

The large discrepancies of OCS intensity and size between E-CON and observations do not change considerably between 2.5 km and 5 km ensembles, meaning that these biases might be associated with unresolved processes (i.e sub-hectometer scales) such as cloud microphysics. For instance, Feng et al. (2018) found that a better representation of stratiform rain results in a better representation of precipitation area in mesoscale convective systems at storm-resolving resolutions. This suggests that microphysical processes might be important for properly representing some macrophysical properties of OCS in the Amazon (e.g. size).

4.2 Diurnal cycle

Considering precipitation only from OCS improves the similarity of the spatial structure in the phase of the diurnal cycle between E-CON and observations (Fig. 7, a, b and c). Especially in the western Amazon, precipitation peaks occurring during night and early morning are as apparent in the 5 km E-CON ensemble as in CMORPH. This feature is noisier in the 2.5 km ensemble probably due to the smaller sample size than the 5 km ensemble.

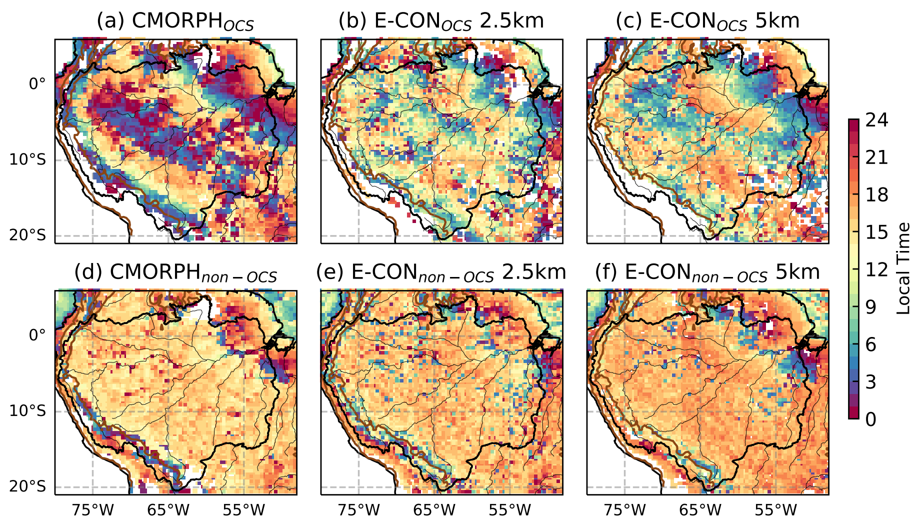

Figure 7. Local time (hour) of maximum precipitation (a, b, c) considering only organized convective systems (OCS) and (d, e, f) only non-organized precipitation (*"non-OCS"*) for (a, d) CMORPH, (b, e) 5 km E-CON and (c, f) 2.5 km E-CON simulations. For the purpose of this figure, the identification of OCS considered regions outside the Amazon. The Amazon basin is defined as black contours as well as the rivers, and the topography at 1000m is shown in brown contours.

The OCS explain most of the spatial heterogeneity in the diurnal cycle of precipitation in observations (compare Fig. 7, a, d and Fig. 4, a). This feature is less in evidence in E-CON over the central Amazon, although the reduced frequency of OCS in the E-CON ensembles (about one third of CMORPH, not shown) may explain the difference with observations. While there is a clear nocturnal maximum in the simulated OCS precipitation, it is delayed by a few hours as compared to observations. CMORPH displays a nocturnal peak preferably between midnight and 3LT, whereas peaks between 3LT to 6LT are more apparent in the 5 km E-CON ensemble. In contrast, diurnal peaks (12LT to 18LT) are more similar between CMORPH and E-CON, especially at 2.5 km. Other features associated with the diurnal cycle of OCS by E-CON are also consistent with the observations. For instance, the largest and less intense OCS are shown in the early morning, consistent with a decay stage of these systems (Houze Jr, 2004); whereas the most intense and smaller OCS take place in the late afternoon in agreement with their mature phase (not shown).

Non-organized precipitation features daytime precipitation maximum ranging mostly from 12 h to 18 h (Fig. 7 d, e, f), with predominantly peaking times at about 15LT in observations, at 16LT in the 2.5 km ensemble and at 18LT in the 5 km E-CON. Despite the overall diurnal peaks some regions display maximum precipitation overnight in both observations and E-CON ensembles. For instance, scattered nocturnal peaks in the central Amazon, probably associated with very intense rain rates from isolated convective cells (Fig.6, a), are placed near the Amazon river and its tributaries .

While diurnal precipitation peaks seem to improve with increased resolution (Sato et al., 2009), especially for non-organized precipitation, nocturnal precipitation peaks associated with OCS remain similar between 2.5 km and 5 km ensembles. This insensitivity to spatial resolution might indicate once more that other unresolved processes are important for representing the correct lifecyle of OCS in the Amazon.

5 Environmental conditions related to OCS

5.1 Classification of OCS

To better understand the structure of the simulated OCS and environmental factors that influence them, we first apply the k-means clustering technique to objectively identify the main types of OCS in terms of their time of occurrence, size, intensity and location (defined as the center of gravity of each OCS) within the whole Amazon. We also use the Silouhette score (Rousseeuw, 1987), which finds the optimum number of clusters based on a measure of cluster cohesion and separation. The analysis is focused on the 5 km E-CON ensemble, whereas CMORPH observations serve only for comparison of the OCS classification.

Six OCS clusters are identified in both E-CON and CMORPH (Table 2, Fig. S1). Among these, a clear distinction is associated with their time of occurrence rather than their size or intensity. Given that OCS represent mature systems, we refer to those that occur in the afternoon (12-18 hrs) as diurnal (D1, D2, D3), and to those that occur in early morning (5-10 hrs) as nocturnal (N1, N2, N3) OCS. Each of them accounts for about 50 % of the total OCS in E-CON (49.4 % for diurnal and 50.6 % for nocturnal); whereas in the observations, diurnal OCSs are more clearly favored (58.4% versus 41.6 %).

Table 2. Summary of clusters features in the 5 km E-CON ensemble and CMORPH (in parentheses). The median values are presented for the local hour, intensity, area and location (latitude and longitude). The last column indicates the fraction that a given cluster represents from the total OCS.

Cluster	Local hour	Intensity ($mm\,h^{-1}$)	Area (km^2)	Latitude	Longitude	Fraction (%)
D1	17 (18)	9.8 (4.9)	15 513.2 (21 742.6)	-2.8 (-9.0)	-64.1 (-60.6)	20.8 (22.9)
D2	12 (13)	9.0 (4.7)	15 561.69 (22 901.6)	-12.3 (-2.9)	-65.5 (-72.0)	16.3 (25.5)
D3	13 (13)	17.3 (8.8)	15 048.7 (25 949.4)	-5.7 (-5.2)	-65.9 (-67.1)	12.3 (10.0)
N1	6 (7)	9.0 (4.8)	16 671.1 (24 898.3)	-3.7 (-3.2)	-58.1 (-58.4)	22.6 (17.8)
N2	8 (5)	8.6 (4.8)	16 793.1 (23 447.9)	-3.4 (-11.5)	-72.6 (-65.5)	22.6 (17.5)
N3	9 (10)	9.4 (5.7)	47 832.2 (102 669.8)	-5.8 (-5.6)	-67.4 (-66.3)	5.4 (6.4)

Other common features between E-CON and CMORPH OCS are found in clusters N1, N3 and D3. The N1-OCS distinguish from other nocturnal OCS because of their center of gravity is placed in the northwest Amazon, which would correspond to the well-known squall lines propagating from the coast (e.g., Garstang et al., 1994). N3 and D3 OCS do not show a preferred location of occurrence but they are characterized by their large size and high intensity, respectively.

Contrasting the remaining OCS (N2, D1 and D2) between E-CON and CMORPH, these mainly differ in their center of gravity and are more symmetrically distributed in the simulations than observations (Fig. S1). For instance, D1-OCS and D2-OCS comprise the northern and southern Amazon in the E-CON ensemble, respectively; whereas

D2-OCS only cover the northwestern Amazon and D1-OCS, the rest of the basin in CMORPH. As opposed to N1-OCS, N2-OCS comprise the northwestern Amazon in E-CON; but they cover a large region in the southern Amazon in CMORPH. Notwithstanding these differences and the overall discrepancies regarding size and intensity of OCS as described in Section 4.1, we conclude that the E-CON ensemble makes a fair representation of the observed classification of OCS.

5.2 Influence of the environment on OCS evolution

We further explore the mean environmental conditions associated with OCS during their evolution by analyzing composites at different lead and lag times. To isolate diurnal and nocturnal events we consider OCS identified from 12LT (for diurnal OCS) and before 10LT (for nocturnal OCS).

Diurnal and nocturnal OCS show clear distinctions in their vertical structure (Fig. 8), with some variations among clusters. For instance, diurnal OCS persist less than the nocturnal OCS at the place of detection. Both the cloud content and vertical velocity are considerably reduced at 3-hour lead in the diurnal clusters (Fig. 8, dotted lines), whereas nocturnal OCS show larger cloud content and vertical velocity from the freezing level (500 hPa) compared to the lower troposphere. This vertical structure suggests persistent and less intense precipitation in the nocturnal OCS consistent with stratiform features, an essential component of mature OCS (Houze Jr, 2004). In contrast, diurnal OCS display enhanced convective activity (i.e vertical ascent) only at the time of OCS detection, which can be related to a shorter life span or faster propagation than the nocturnal OCS.

The diurnal OCS are associated with a strong (2 K) depression of the surface potential temperature relative to the environment (Fig. 9). This local signal is thought to be related to cold pools as shown in Figure 10. The time of detection of cold pools displays successive times of occurrence from 12LT to 17LT in agreement with the OCS propagation, especially in the northeastern Amazon. Figure 9 shows westward-propagating anomalies of potential temperature which are stronger 3 hours before than after the detection of OCS, in agreement with the occurrence of cold pools mainly during the early afternoon. The westward propagation is consistent with the background zonal flow (Fig. 8), which displays easterlies through a deep layer (from 950 hPa to 400 hPa). This propagation is more evident in D1-OCS (northern Amazon) and D3-OCS (most intense OCS),

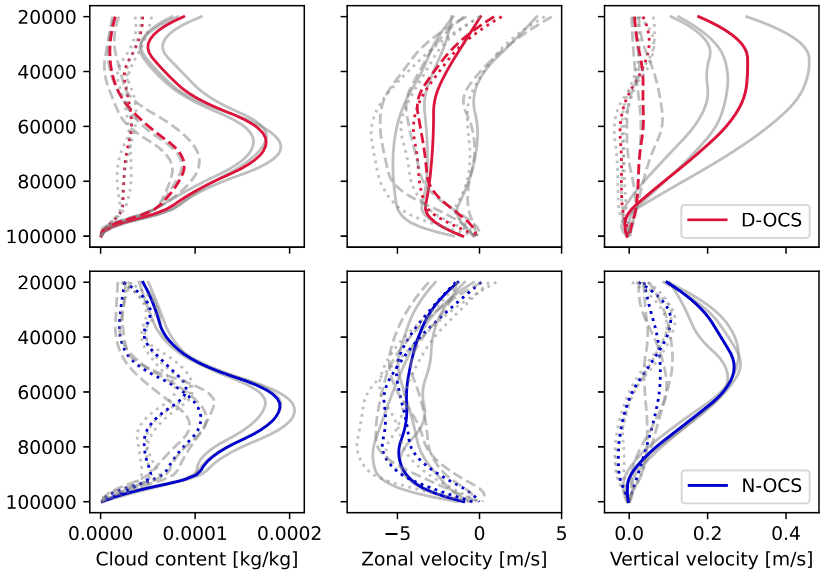

Figure 8. Vertical profile of composite OCS. The variables shown are from left to right: cloud content (water and ice), zonal velocity and vertical velocity. Solid contours represent the vertical profiles at the moment of object detection (time 0), dashed lines represent 3 hours before the detection and dotted lines, 3 hours after time 0. Diurnal (D-OCS) and nocturnal (N-OCS) OCS are located in the upper and lower row, respectively. Grey contours represent the original clusters (D1, D2, D3, N1, N2, N3). The vertical profiles are smoothed using a second order polinomial interpolation.

whereas D2-OCS (southern Amazon) show rather a stationary pattern, probably due to their far distance from the trade winds (not shown).

The nocturnal OCS are associated with different large scale conditions as compared to the diurnal OCS. The zonal wind velocity is considerably larger near 800 hPa (Fig. 8) than the surface even 3 hours after the OCS detection, especially for N1-OCS (northeastern Amazon) and N3-OCS (largest OCS). The strong easterlies in the lower troposphere can be indicative of the nocturnal low-level jet (Anselmo et al., 2020), which would act against the stable nocturnal boundary layer to sustain convection overnight (e.g., Houze Jr, 2004). Anselmo et al. (2020) found enhanced occurrence of cloud clusters associated with such nocturnal low-level jet, especially during the early morning (2LT to 8LT) which is in agreement with our results. Moreover, the potential temperature perturbations show larger anomalies above the surface (850 hPa) during the OCS occurrence (Fig. 9). A spe-

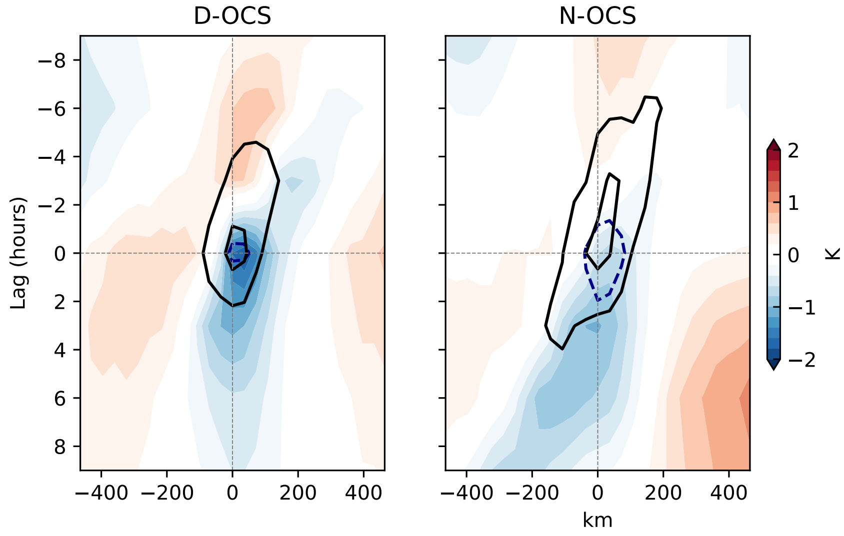

Figure 9. Time-longitude composites of potential temperature perturbations at the surface (1000hPa) related to diurnal (left) and nocturnal (right) OCS. Negative anomalies at 850hPa are shown as dashed-blue contours ($-0.3\,\mathrm{K}$). The anomalies are computed with respect to the zonal mean. Precipitation is displayed as black solid contours (0.2 and $0.5\,\mathrm{mm\,h^{-1}}$). Time zero indicates the hour when the objects are detected and longitude zero is the location of the center of mass of the precipitating objects.

cial case is noted in N3-OCS (not shown), which display broader anomalies of potential temperature that last 3 hours after their detection. These elevated anomalies might be related to cooling by evaporation of precipitation particles and indicative of their decay stage.

The environmental controls of diurnal and nocturnal OCS as represented by explicitly resolved convection, provide insights of which processes might be important and could be improved. For instance, it was mentioned that the frequency of OCS is considerably less in the E-CON ensembles than observations. More precisely, it appears that the diurnal OCS are those rather underestimated. The results show that surface processes matter mainly for the diurnal OCS, which suggest that these processes in relation to deep precipitating convection might need to be better represented at storm-resolving scales.

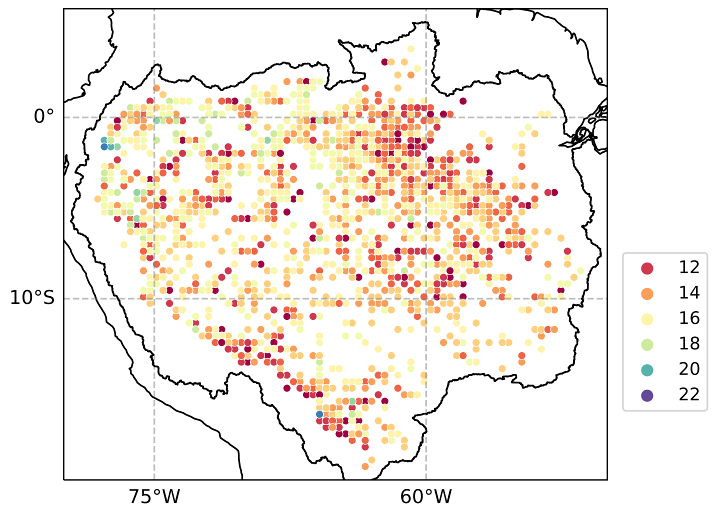

Figure 10. Local hour of cold pool detection during the time of occurrence of OCS. Cold pools are identified considering a potential temperature perturbation larger than $-2\,\mathrm{K}$ and precipitation greater than $1\,\mathrm{mm\,h^{-1}}$.

6 Summary and conclusion

This study investigates the ability of storm-resolving simulations to represent precipitating systems over the Amazon river basin. We perform ensemble simulations with the ICON-NWP atmospheric model at a coarse grid spacing (40 km) wherein convection is parameterized (P-CON) and storm-resolving simulations that enable the explicit representation of convection (E-CON) at 2.5 km and 5 km grid spacing. The simulations are compared to each other, conventional coarse resolution model output taken from CMIP, and to observations as represented by the CMPORPH dataset.

The mean precipitation in the Amazon basin and its spatial distribution is fairly represented by both E-CON and P-CON ensembles. However, the large frequency of light rain can explain a close daily mean to observations in the P-CON ensemble. Moreover, P-CON misses precipitation in the northeast coast which is known to be important for the generation of propagating systems towards the Amazon (e.g., Greco et al., 1990; Burleyson et al., 2016; Rehbein et al., 2018).

Ensembles with grid spacings that allow for explicit convection better represent the distribution of precipitation intensity and the spatial variability of the diurnal cycle, as compared to simulations with parameterized convection. Light-to-moderate precipitation $2\,\mathrm{mm\,d^{-1}}$ to $20\,\mathrm{mm\,d^{-1}}$ and higher intensity rain rates are correctly captured by E-CON; whereas P-CON persists on long-standing biases (e.g., Stephens et al., 2010) as the CMIP models. The spatial heterogeneity (pattern) of the diurnal cycle can also be detected in the E-CON ensemble, similar to what is found in observations. The P-CON ensemble, which is based on one of the best tested and most advanced parameterization schemes, while able to reproduce the afternoon peak of maximum precipitation over most of the Amazon in contrast to the CMIP models, its spatial distribution is rather homogeneous and misses the nocturnal precipitation over the central and northeast Amazon.

The E-CON ensemble shows evidence of organized convective systems that are absent in the P-CON ensemble. These OCS are shown to be closely associated with the better representation of Amazon precipitation, as they explain the frequency of high intense rain rates and the heterogeneity of the precipitation diurnal cycle in observations. The similarity between E-CON and observations improves in both the distribution of precipitation intensity and diurnal cycle when only considering precipitation from OCS. However, the simulated OCS simulated by the E-CON ensemble are still less frequent, smaller and more intense than observed.

The simulated and observed OCS cluster into nocturnal and diurnal systems. The environment of the nocturnal versus diurnal systems differs systematically. Nocturnal clusters are associated with stronger easterlies in the lower troposphere, peaking at about 850 hPa and forming part of the Amazonian low-level jet (Anselmo et al., 2020). In addition, an elevated cooler atmosphere propagates with the OCS during the early morning. Not surprisingly, the diurnal OCS show a stronger signature at the surface than the nocturnal OCS. The E-CON simulations suggest that cold pools contribute to the propagation of OCS in the northern Amazon and those that are very intense (D3-OCS). A composite analysis over diurnal clusters shows a strong temperature perturbation at the surface that propagates during the early afternoon. Given that the simulations produced about 20 % less of diurnal OCS than observations, such systems may be sensitive to the representation of surface processes in ways that the E-CON simulations insufficiently capture.

Our simulations show a clear improvement in many aspects of precipitation over the Amazon river basin when the precipitating systems are simulated explicitly. By simulating the geometry and transient dynamics of precipitating convection systems, a better representation of organized convective systems emerge, and these prove essential for capturing many features of the observed precipitation over the Amazon. Nonetheless, our simulations also show room for improvement. For instance in representing the relative prominence of organized systems during the day, which may be sensitive to land-surface processes, and in the timing of the nocturnal peak of precipitation, which has previously been shown to be sensitive to cloud microphysical processes (e.g., Feng et al., 2018). Simulations with a twofold finer grid do not lead to dramatic improvements, indicating that to the extent deficiencies are related to a poor representation of small scale circulations, capturing these effects explicitly would require much (hecto to deca meter) finer resolution.

7 Open Research

HadISST data (Rayner et al., 2003) is available at https://www.metoffice.gov.uk/hadobs/hadisst/data/down CMORPH precipitation dataset (Xie et al., 2017) were obtained from https://www.ncei.noaa.gov/data/cmorph-high-resolution-global-precipitation-estimates/access/30min/8km. Boundaries of the Amazon are available at Lehner et al. (2006) and were obtained from http://hydrosheds.cr.usgs.gov. CMIP output of this study were replicated and made available for this study by the German Climate Computing Centre (Deutschen Klimarechenzentrum, DKRZ) and pre-processed at Fiedler et al. (2020). Primary data and scripts used in the analysis that may be useful in reproducing the author's work are archived by the Max Planck Institute for Meteorology and can be obtained via the institutional repository https://pure.mpg.de

Acknowledgments

This study was supported by the Max Planck Society for the Advancement of Science and . The authors thank Jaemyeong Seo for providing scripts for the cold pool detection and Cathy Hohenegger for her useful and constructive comments on the paper.

References

Anselmo, E. M., Machado, L. A., Schumacher, C., & Kiladis, G. N. (2021). Amazonian mesoscale convective systems: Life cycle and propagation characteristics. *International Journal of Climatology*.

Anselmo, E. M., Schumacher, C., & Machado, L. A. (2020). The amazonian low-level jet and its connection to convective cloud propagation and evolution.

Monthly Weather Review, *148*(10), 4083–4099.

Arnold, N. P., Putman, W. M., & Freitas, S. R. (2020). Impact of resolution and parameterized convection on the diurnal cycle of precipitation in a global nonhydrostatic model. *Journal of the Meteorological Society of Japan. Ser. II*.

Bechtold, P. (2017). Atmospheric moist convection. *Meteorological Training Course Lecture Series*, 1-78. Retrieved from `https://www.ecmwf.int/node/16953`

Bechtold, P., Köhler, M., Jung, T., Doblas-Reyes, F., Leutbecher, M., Rodwell, M. J., ... Balsamo, G. (2008). Advances in simulating atmospheric variability with the ecmwf model: From synoptic to decadal time-scales. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, *134*(634), 1337–1351.

Becker, T., Bechtold, P., & Sandu, I. (2021). Characteristics of convective precipitation over tropical Africa in storm-resolving global simulations. *Quarterly Journal of the Royal Meteorological Society*, *147*(741), 4388–4407.

Betts, A. K., & Jakob, C. (2002). Study of diurnal cycle of convective precipitation over Amazonia using a single column model. *Journal of Geophysical Research: Atmospheres*, *107*(D23), ACL–25.

Birch, C. E., Roberts, M. J., Garcia-Carreras, L., Ackerley, D., Reeder, M. J., Lock, A. P., & Schiemann, R. (2015). Sea-breeze dynamics and convection initiation: The influence of convective parameterization in weather and climate model biases. *Journal of Climate*, *28*(20), 8093–8108.

Burleyson, C. D., Feng, Z., Hagos, S. M., Fast, J., Machado, L. A., & Martin, S. T. (2016). Spatial variability of the background diurnal cycle of deep convection around the GoAmazon2014/5 field campaign sites. *Journal of Applied Meteorology and Climatology*, *55*(7), 1579–1598.

Chavez, S. P., & Takahashi, K. (2017). Orographic rainfall hot spots in the Andes-Amazon transition according to the TRMM precipitation radar and in situ data. *Journal of Geophysical Research: Atmospheres*, *122*(11), 5870–5882.

Crook, J., Klein, C., Folwell, S., Taylor, C. M., Parker, D. J., Stratton, R., & Stein, T. (2019). Assessment of the representation of West African storm lifecycles in convection-permitting simulations. *Earth and Space Science*, *6*(5), 818–835.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project

Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958.

Feng, Z., Leung, L. R., Houze Jr, R. A., Hagos, S., Hardin, J., Yang, Q., . . . Fan, J. (2018). Structure and evolution of mesoscale convective systems: Sensitivity to cloud microphysics in convection-permitting simulations over the united states. *Journal of Advances in Modeling Earth Systems*, *10*(7), 1470–1494.

Feng, Z., Leung, L. R., Liu, N., Wang, J., Houze Jr, R. A., Li, J., . . . Guo, J. (2021). A Global High-resolution Mesoscale Convective System Database Using Satellite-Derived Cloud tops, Surface Precipitation, and Tracking. *Journal of Geophysical Research: Atmospheres*, *126*(8), e2020JD034202.

Fiedler, S., Crueger, T., D'Agostino, R., Peters, K., Becker, T., Leutwyler, D., . . . others (2020). Simulated tropical precipitation assessed across three major phases of the Coupled Model Intercomparison Project (CMIP). *Monthly Weather Review*, *148*(9), 3653–3680.

Fitzjarrald, D. R., Sakai, R. K., Moraes, O. L., de Oliveira, R. C., Acevedo, O. C., Czikowsky, M. J., & Beldini, T. (2008). Spatial and temporal rainfall variability near the Amazon-tapajós confluence. *Journal of Geophysical Research: Biogeosciences*, *113*(G1).

Garreaud, R., & Wallace, J. M. (1997). The diurnal march of convective cloudiness over the Americas. *Monthly Weather Review*, *125*(12), 3157–3171.

Garstang, M., Massie Jr, H. L., Halverson, J., Greco, S., & Scala, J. (1994). Amazon coastal squall lines. Part I: Structure and kinematics. *Monthly Weather Review*, *122*(4), 608–622.

Greco, S., Swap, R., Garstang, M., Ulanski, S., Shipham, M., Harriss, R., . . . Artaxo, P. (1990). Rainfall and surface kinematic conditions over central Amazonia during ABLE 2B. *Journal of Geophysical Research: Atmospheres*, *95*(D10), 17001–17014.

Heise, E., Ritter, B., Schrodin, R., & Wetterdienst, D. (2006). *Operational implementation of the multilayer soil model.* Citeseer.

Herbert, R., Stier, P., & Dagan, G. (2021). Isolating large-scale smoke impacts on cloud and precipitation processes over the amazon with convection permitting resolution. *Journal of Geophysical Research: Atmospheres*, *126*(13).

Holloway, C., Woolnough, S., & Lister, G. (2012). Precipitation distributions for explicit versus parametrized convection in a large-domain high-resolution tropical case study. *Quarterly Journal of the Royal Meteorological Society*, *138*(668), 1692–1708.

Houze Jr, R. A. (2004). Mesoscale convective systems. *Reviews of Geophysics*, *42*(4).

Houze Jr, R. A., Rasmussen, K. L., Zuluaga, M. D., & Brodzik, S. R. (2015). The variable nature of convection in the tropics and subtropics: A legacy of 16 years of the tropical rainfall measuring mission satellite. *Reviews of Geophysics*, *53*(3), 994–1021.

Inoue, T., Rajendran, K., Satoh, M., & Miura, H. (2021). On the semidiurnal variation in surface rainfall rate over the tropics in a global cloud-rresolving model simulation and satellite observations. *Journal of the Meteorological Society of Japan. Ser. II*.

Janowiak, J. E., Kousky, V. E., & Joyce, R. J. (2005). Diurnal cycle of precipitation determined from the cmorph high spatial and temporal resolution global precipitation analyses. *Journal of Geophysical Research: Atmospheres*, *110*(D23).

Judt, F., & Rios-Berrios, R. (2021). Resolved convection improves the representation of equatorial waves and tropical rainfall variability in a global nonhydrostatic model. *Geophysical Research Letters*.

Lehner, B., Verdin, K., & Jarvis, A. (2006). Hydrological data and maps based on Shuttle elevation derivatives at multiple scales (HydroSHEDS)-technical documentation. *World Wildlife Fund US, Washington, DC*.

Love, B. S., Matthews, A. J., & Lister, G. M. (2011). The diurnal cycle of precipitation over the maritime continent in a high-resolution atmospheric model. *Quarterly Journal of the Royal Meteorological Society*, *137*(657), 934–947.

Mapes, B., & Neale, R. (2011). Parameterizing convective organization to escape the entrainment dilemma. *Journal of Advances in Modeling Earth Systems*, *3*(2).

Marengo, J. A. (2006). On the hydrological cycle of the amazon basin: A historical review and current state-of-the-art. *Revista brasileira de meteorologia*, *21*(3), 1–19.

Paccini, L., Hohenegger, C., & Stevens, B. (2021). Explicit versus parameterized convection in response to the Atlantic meridional mode. *Journal of Climate*, *34*(9), 3343–3354.

Pereira Filho, A. J., Carbone, R. E., Tuttle, J. D., Karam, H. A., et al. (2015). Convective rainfall in Amazonia and adjacent tropics. *Atmospheric and Climate Sciences*, *5*(02), 137.

Phillips, O. L., Lewis, S. L., Baker, T. R., Chao, K.-J., & Higuchi, N. (2008). The changing amazon forest. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1498), 1819–1827.

Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., … others (2015). A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Reviews of geophysics*, *53*(2), 323–361.

Rayner, N., Parker, D. E., Horton, E., Folland, C. K., Alexander, L. V., Rowell, D., … Kaplan, A. (2003). Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research: Atmospheres*, *108*(D14).

Rehbein, A., Ambrizzi, T., & Mechoso, C. R. (2018). Mesoscale convective systems over the Amazon basin. Part I: climatological aspects. *International Journal of Climatology*, *38*(1), 215–229.

Richter, I., & Xie, S.-P. (2008). On the origin of equatorial atlantic biases in coupled general circulation models. *Climate Dynamics*, *31*(5), 587–598.

Rickenbach, T. M. (2004). Nocturnal cloud systems and the diurnal variation of clouds and rainfall in southwestern amazonia. *Monthly Weather Review*, *132*(5), 1201–1219.

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, *20*, 53–65.

Santos, M. J., Medvigy, D., Silva Dias, M. A., Freitas, E. D., & Kim, H. (2019). Seasonal flooding causes intensification of the river breeze in the central amazon. *Journal of Geophysical Research: Atmospheres*, *124*(10), 5178–5197.

Sato, T., Miura, H., Satoh, M., Takayabu, Y. N., & Wang, Y. (2009). Diurnal cycle of precipitation in the tropics simulated in a global cloud-resolving model. *Journal of Climate*, *22*(18), 4809–4826.

Satoh, M., Stevens, B., Judt, F., Khairoutdinov, M., Lin, S.-J., Putman, W. M., & Düben, P. (2019). Global cloud-resolving models. *Current Climate Change*

Reports, *5*(3), 172–184.

Slingo, J., Bates, P., Bauer, P., Belcher, S., Palmer, T., Stephens, G., ... Teutsch, G. (2022). Ambitious partnership needed for reliable climate prediction. *Nature Climate Change*, *12*(6), 499–503.

Stephens, G. L., L'Ecuyer, T., Forbes, R., Gettelmen, A., Golaz, J.-C., Bodas-Salcedo, A., ... Haynes, J. (2010). Dreary state of precipitation in global models. *Journal of Geophysical Research: Atmospheres*, *115*(D24).

Stevens, B., Acquistapace, C., Hansen, A., Heinze, R., Klinger, C., Klocke, D., ... others (2020). The added value of large-eddy and storm-resolving models for simulating clouds and precipitation. *Journal of the Meteorological Society of Japan. Ser. II*.

Tai, S.-L., Feng, Z., Ma, P.-L., Schumacher, C., & Fast, J. D. (2021). Representations of precipitation diurnal cycle in the amazon as simulated by observationally constrained cloud-system resolving and global climate models. *Journal of Advances in Modeling Earth Systems*, *13*(11), e2021MS002586.

Tanaka, L. d. S., Satyamurty, P., & Machado, L. A. T. (2014). Diurnal variation of precipitation in central Amazon Basin. *International Journal of Climatology*, *34*(13), 3574–3584.

Tang, S., Gleckler, P., Xie, S., Lee, J., Ahn, M.-S., Covey, C., & Zhang, C. (2021). Evaluating the diurnal and semidiurnal cycle of precipitation in CMIP6 Models Using Satellite-and Ground-Based Observations. *Journal of Climate*, *34*(8), 3189–3210.

Wu, M., Lee, J.-E., Wang, D., & Salameh, M. (2021). Suppressed Daytime Convection over the Amazon River. *Journal of Geophysical Research: Atmospheres*, e2020JD033627.

Xie, P., Joyce, R., Wu, S., Yoo, S.-H., Yarosh, Y., Sun, F., & Lin, R. (2017). Reprocessed, bias-corrected CMORPH global high-resolution precipitation estimates from 1998. *Journal of Hydrometeorology*, *18*(6), 1617–1641.

Yang, S., Kuo, K.-S., & Smith, E. A. (2008). Persistent nature of secondary diurnal modes of precipitation over oceanic and continental regimes. *Journal of Climate*, *21*(16), 4115–4131.

Yin, L., Fu, R., Shevliakova, E., & Dickinson, R. E. (2013). How well can cmip5 simulate precipitation and its controlling processes over tropical south america? *Climate Dynamics*, *41*(11-12), 3127–3143.

Zängl, G., Reinert, D., Rípodas, P., & Baldauf, M. (2015). The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core. *Quarterly Journal of the Royal Meteorological Society*, *141*(687), 563–579.