Geoscientific
Model Development

Model intercomparison of COSMO 5.0 and IFS 45r1 at kilometer-scale grid spacing

Christian Zeman[1], **Nils P. Wedi**[2], **Peter D. Dueben**[2], **Nikolina Ban**[3], and **Christoph Schär**[1]

[1]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland
[2]European Centre for Medium-Range Weather Forecasts, Reading, UK
[3]Department of Atmospheric and Cryospheric Sciences, University of Innsbruck, Innsbruck, Austria

Correspondence: Christian Zeman (christian.zeman@env.ethz.ch)

Abstract. The increase in computing power and recent model developments allow for the use of global kilometer-scale weather and climate models for routine forecasts. At these scales, deep convective processes can be partially resolved explicitly by the model dynamics. Next to horizontal resolution, other aspects such as the applied numerical methods, the use of the hydrostatic approximation, and time step size are factors that might influence a model's ability to resolve deep convective processes.

In order to improve our understanding of the role of these factors, a model intercomparison between the nonhydrostatic COSMO model and the hydrostatic Integrated Forecast System (IFS) from ECMWF has been conducted. Both models have been run with different spatial and temporal resolutions in order to simulate 2 summer days over Europe with strong convection. The results are analyzed with a focus on vertical wind speed and precipitation.

Results show that even at around 3 km horizontal grid spacing the effect of the hydrostatic approximation seems to be negligible. However, time step proves to be an important factor for deep convective processes, with a reduced time step generally allowing for higher updraft velocities and thus more energy in vertical velocity spectra, in particular for shorter wavelengths. A shorter time step is also causing an earlier onset and peak of the diurnal cycle. Furthermore, the amount of horizontal diffusion plays a crucial role for deep convection with more diffusion generally leading to larger convective cells and higher precipitation intensities. The study also shows that for both models the parameterization of deep convection leads to lower updraft and pre-cipitation intensities and biases in the diurnal cycle with a precipitation peak which is too early.

1 Introduction

The Earth's atmosphere is home to processes ranging from scales as large as the planet itself, such as the trade winds, down to scales of angstroms (10^{-10} m), such as Rayleigh scattering of sunlight by an air molecule. Explicitly resolving all these processes in an atmospheric model is virtually impossible, even in the distant future. But the ever greater availability of computing power allows us to at least come closer by reducing spatial resolutions in numerical weather prediction and climate models step by step (Schulthess et al., 2019; Neumann et al., 2019). One of the processes that can nowadays be resolved is deep convection: the rise of buoyant plumes, strong enough to break through the temperature inversions and reaching as high as to the tropopause. Given that there is enough moisture in the air, the plumes can form towering cumulonimbus clouds and cause heavy thunderstorms. On a larger scale, deep convection is an important process for the redistribution of heat, moisture, and momentum with a subsequent large impact on the general circulation in the atmosphere (Houze and Betts, 1981; Held and Soden, 2006).

Atmospheric models with a grid spacing of around 4 km and smaller have been considered to at least partially resolve deep convection (Weisman et al., 1997; Romero et al., 2001; Done et al., 2004), while models with coarser resolutions have to rely on parameterization of deep convection. One of the drawbacks of parameterized deep convection is a known

phase error in the diurnal cycle of precipitation, by being too closely coupled to the phase of solar radiation and thus peaking too early (Yang and Slingo, 2001; Betts and Jakob, 2002; Guichard et al., 2004; Dai and Trenberth, 2004), even though more recent developments by Bechtold et al. (2014) have shown some improvements in this regard. Coarse models with parameterized deep convection also tend to overestimate precipitation frequency but underestimate precipitation intensity (Dai and Trenberth, 2004; Sun et al., 2006; Stephens et al., 2010). The explicit treatment of deep convection usually leads to a better representation of the diurnal cycle (Hohenegger et al., 2008; Dirmeyer et al., 2012; Ban et al., 2014; Pearson et al., 2014; Leutwyler et al., 2017), a better spatial representation of rainfall (Kendon et al., 2012; Prein et al., 2013), and more realistic hourly intensities of extreme precipitation events (Prein et al., 2013; Ban et al., 2014; Fosser et al., 2014; Kendon et al., 2019). More details about the benefits of kilometer-scale climate models can be found in several review articles (Prein et al., 2015; Schär et al., 2020).

Nonetheless, deep convection is not yet fully resolved with grid spacings of 1–4 km. To fully resolve deep convection, one would require a grid spacing of around 250 m or less (Bryan et al., 2003; Lebo and Morrison, 2015; Jeevanjee, 2017). But even though the structural convergence of updrafts and clouds is not yet reached at a kilometer scale, many domain-averaged and integrated properties related to a large ensemble of convective cells (i.e., mean diurnal cycle, spatial distribution of precipitation, clouds, diabatic heating, convective transport of mass, heat, and water vapor) have been shown to converge already at a grid spacing of around 4 km (Panosetti et al., 2018, 2019). This so-called bulk convergence makes the explicit treatment of deep convection at the kilometer scale an attractive practice, as it can bring the aforementioned improvements without paying the huge computational costs associated with fully resolving deep convection. Recent work by Vergara-Temprado et al. (2020) has shown that the explicit treatment of deep convection may already be beneficial even at relatively coarse grid spacings of up to 25 km for selected metrics such as hourly precipitation statistics and the representation of the diurnal cycle over nonorographic regions.

Another subject that is often associated with weather and climate models at the kilometer scale is the use of the hydrostatic approximation in the governing equations. The hydrostatic approximation assumes the vertical accelerations to be small compared to the buoyancy force. This is normally the case when the horizontal length scale of the flow is much larger than the vertical length scale. With the hydrostatic approximation, vertical velocity can be derived from the continuity equation and thus becomes a diagnostic variable. The resulting system of equations is simpler and usually computationally less expensive to solve, which makes it an attractive option for models as long as the hydrostatic approximation is still suitable. For example, the nonhydro-

static version of the Integrated Forecast System (IFS) model from the European Centre for Medium-Range Weather Forecasts (ECMWF) is about 80 % more expensive than the corresponding hydrostatic version at a grid spacing of around 9 km (Wedi et al., 2009).

There is not really a consensus in the scientific community about the horizontal resolution at which the hydrostatic approximation is no longer suitable. For example, Ross and Orlanski (1978) performed a two-dimensional simulation of an idealized cold front and found no big differences between the hydrostatic and the nonhydrostatic setup for a grid spacing of 20 km, while Orlanski (1981) found significant differences for a similar case with 8 km grid spacing. According to calculations by Daley (1988), models with a grid spacing of 25 km or smaller should already use the nonhydrostatic set of equations. But then again, Dudhia (1993) found only little differences between the hydrostatic and the nonhydrostatic solution for a cold front with grid spacing of 6.5 km. Kato and Saito (1995) performed idealized moist convection simulations with grid spacings of 20, 10, and 5 km and found that the hydrostatic model without parameterized deep convection overdevelops updrafts and overestimates convective precipitation amount and area. These results were later confirmed for a real-world case from Kato (1996). Kato (1997) recommends the use of moist convective parameterization (e.g., Manabe et al., 1965) when using a hydrostatic model with around 10 km grid spacing and the use of a nonhydrostatic model for a grid spacing of 5 km. Recent global realworld simulations with the hydrostatic IFS at a grid spacing of 1.45 km by Dueben et al. (2020) and Wedi et al. (2020) produced realistic results and did not show deficiencies that could be directly attributed to the invalidity of the hydrostatic assumption at this resolution.

Several studies also primarily looked at the vertical velocities of hydrostatic and nonhydrostatic models at different resolutions. A maybe counterintuitive behavior of the hydrostatic regime is the development of too high vertical wind velocities at resolutions where the hydrostatic assumption is no longer valid. This is due to the fact that the vertical wind velocity is directly diagnosed from the horizontal velocities, and there is no nonhydrostatic process limiting the vertical mass flux. Simulations of a squall line with horizontal grid spacings reaching from 20 to 1 km by Weisman et al. (1997) showed the hydrostatic model overestimating the maximum vertical velocity at grid spacings of 4 km and lower. Jeevanjee (2017) found an overestimation of vertical velocities of the hydrostatic model at grid spacings smaller than 2 km in radiative-convective-equilibrium simulations over sea with grid spacings ranging from 16 to 0.0625 km. Dueben et al. (2020) performed global simulations with IFS using the hydrostatic and nonhydrostatic equations at 1.45 km grid spacing where the updraft velocities were quite similar when using a time step of 30 s.

Not only the system of equations but also the applied numerical methods are important when it comes to under-

standing the model behavior. The two models used for this study are very different in this regard: while the hydrostatic IFS model is a spectral model with a semi-Lagrangian semi-implicit scheme, the nonhydrostatic COSMO model is a Eulerian model with a split explicit scheme in the horizontal and a implicit scheme in the vertical dimension. These differences in design have direct implications on the conditions for numerical stability and the associated time step of the models. Thanks to the semi-Lagrangian treatment of advection, the time step in IFS is not limited by the Courant–Friedrichs–Lewy (CFL) condition (Courant et al., 1928) but by the Lipschitz condition. The Lipschitz condition requires the time step to be smaller than the reciprocal of the absolute maximum value of the wind shear at each direction (Pudykiewicz et al., 1985; Staniforth and Côté, 1991). It ensures that the trajectories do not intersect each other (Smolarkiewicz and Pudykiewicz, 1992) and is less restrictive regarding time step than the CFL condition, allowing an atmospheric model to remain stable and deliver accurate results even with CFL numbers higher than 4 (Staniforth and Côté, 1991). This implies that IFS can be run with a rather long time step and still remain stable, even though the CFL condition might be violated at some locations with high wind speed. In contrast, COSMO uses an Eulerian explicit approach for horizontal advection and thus the time step has to be small enough to not violate the horizontal CFL condition at any location in order to guarantee stability. Evidence that the semi-Lagrangian semi-implicit scheme does not only provide numerical stability for such high CFL numbers but also produces reasonable results in real-world scenarios is provided by the day-to-day forecasts of ECMWF, using a competitive time step of 450 s at 9 km grid with IFS.

Compared to the many studies addressing spatial resolution in atmospheric models, the sensitivity to temporal resolution has received relatively little attention. Several studies identified time step as a very important factor when it comes to precipitation patterns (Williamson and Olson, 2003), precipitation intensity (Mishra and Sahany, 2011), or tropospheric circulation (Jung et al., 2012). However, these studies were carried out with relatively coarse resolution ($\Delta x >$ 100 km) with parameterized deep convection and are therefore hardly comparable to the resolutions used in this study. Fuhrer et al. (2018) recommends a time step smaller than 40–60 s at around 1 to 2 km grid spacing. The global simulations with IFS from Dueben et al. (2020) with a 1.45 km grid spacing showed improvements in the representation of deep convective processes for both the hydrostatic and nonhydrostatic version when reducing the time step from 120 to 30 s. It is probably difficult to give a generally valid recommendation for time step size, as many different processes are affected by it. Next to the dynamics, the parameterization of subgrid-scale processes, its call frequency, and the type of coupling to the model dynamics (see for example Ubbiali et al., 2021) also have to be considered. For instance, Barrett et al. (2019) performed idealized simulations using COSMO with 1 km

grid spacing and found a 53 % reduction in precipitation with a two-moment microphysics scheme when the time step was increased from 1 to 15 s. These changes were attributed to the time step dependence of the amount of supersaturation with respect to liquid in strong updrafts and the corresponding sensitivity of the cloud microphysics parameterization to this value in combination with the sequential-update splitting coupling. In the current study we use bulk microphysics schemes, which have a much smaller sensitivity with respect to the time step.

Deep convection is a dynamical process that is often happening very locally, involving only a few grid points in kilometer-scale models. The dynamics and concentration of moist variables at such scales are largely affected by diffusion. Diffusion may serve many purposes, such as eliminating numerical noise, increasing model stability, absorbing vertically propagating gravity waves at the model top, or also emulating cumulative effects of unresolved subgrid-scale processes (see Jablonowski and Williamson, 2011, for an overview of diffusion). Next to implicit diffusion, which is inherently caused by the numerical methods, most models apply some form of explicit diffusion. A significant amount of diffusion is also caused by subgrid-scale parameterizations and orography filtering (Malardel and Wedi, 2016). All these aspects can lead to very different model behavior in terms of dissipation, which then might again influence deep convection.

Ricard et al. (2013) conducted a case study over southwest France with the nonhydrostatic limited-area model AROME (Météo-France), which uses the same dynamics as the nonhydrostatic version of IFS, at 2.5 km grid spacing in order to determine the influence of horizontal diffusion on convective cells. They have compared the results to the Eulerian research model Meso-NH (Lafore et al., 1998) and found that AROME develops larger convective cells than Meso-NH with a tendency of the cells to structure into circular patterns with too strong outflow at the surface induced by cooling from precipitation evaporation, especially with additional explicit diffusion. In order to prevent too high precipitation intensities and unrealistic divergent winds at the edges of the cold outflow, Malardel and Ricard (2015) introduced a correction of the interpolation weights of the semi-Lagrangian scheme for IFS, AROME, and HARMONIE (HIRLAM consortium), which all use the same dynamics. These new interpolation weights improve the conservation property of the scheme, effectively increase diffusion in the convergent parts of the flow, and reduce diffusion in the divergent parts of the flow. They performed idealized and real-case experiments of convective systems with the new interpolation weights, which lead to a significant reduction of extreme precipitation with the new interpolation weights. While the experiments shown in the paper were performed with the nonhydrostatic version of the dynamics of IFS, the operational hydrostatic version showed the same improvements.

In order to improve our understanding about the role of some of the aforementioned factors in the representation of deep convection, we here present a model intercomparison between COSMO and IFS, addressing the following key questions:

1. What are the main differences between COSMO and IFS in the representation of deep convective precipitation? How do the precipitation patterns, precipitation intensities, and the diurnal cycle of precipitation look like at different resolutions, and how do they compare with observations?

2. Can we detect significant differences due to the use of the hydrostatic and nonhydrostatic dynamics in IFS and COSMO, respectively?

3. What is the effect of the time step on deep convection in the different models? Are there any disadvantages to using a large time step for simulations with explicit treatment of deep convection?

4. How does horizontal diffusion affect deep convection? Do the two models show differences in the representation of deep convection that can be accounted for by differences in horizontal diffusion?

It has to be mentioned that comparing two so fundamentally different and complex models in simulating a real-world case makes it intriguing but also very difficult to confidently attribute any disparities in the examined results to a specific process or its associated handling in the model. Nevertheless, some assumptions can be made based on findings in previous studies and our knowledge of the different model properties. We would also like to emphasize that this study is not intended to be a performance comparison of the two models. The models have not been specifically tuned for the respective resolution setups and, with only 2 d of hourly data, the sample size is too small to draw any conclusions regarding the general forecast quality. Differences in the results are expected, as the specific weather situation under consideration was difficult to predict and the general setup of the models is different. In particular, IFS is initialized from its own analysis and then run globally, while COSMO is driven by the IFS operational analysis. So the aim of the study is mainly to show differences and similarities of these two very distinct modeling approaches with different configurations in their treatment of deep convective processes and through this extend our knowledge on this subject.

2 Methods and data

2.1 Model description

2.1.1 COSMO

The Consortium for Small-scale Modeling (COSMO) model (Baldauf et al., 2011) has been originally developed for numerical weather prediction but was extended to also run in climate mode (Rockel et al., 2008). COSMO is a regional model and operates on a grid with rotated latitude–longitude coordinates. It uses a split explicit third-order Runge–Kutta discretization (Wicker and Skamarock, 2002) in combination with a fifth-order upwind scheme for horizontal advection and an implicit Crank–Nicolson scheme for vertical advection. Parameterizations used in this version include subgrid-scale orography (SSO) by Lott and Miller (1997), a radiation scheme based on the δ-two-stream approach (Ritter and Geleyn, 1992), a single-moment cloud microphysics scheme (Reinhardt and Seifert, 2006), a turbulent kinetic-energy-based parameterization for the planetary boundary layer (Raschendorfer, 2001), an adapted version of the convection scheme by Tiedtke (1989), and a multi-layer soil model with a representation of groundwater (Schlemmer et al., 2018). Explicit horizontal diffusion is applied by using a monotonic fourth-order linear scheme acting on model levels for wind, temperature, pressure, specific humidity, and cloud water content (Doms and Baldauf, 2018). An orographic limiter helps avoid excessive vertical mixing around mountains. For the standard experiments in this paper, the explicit diffusion from the monotonic fourth-order linear scheme is set to zero. However, we apply Smagorinsky diffusion (Smagorinsky, 1963) to the horizontal wind components for all experiments in order to enhance the numerical stability of the scheme in the presence of horizontal shear instabilities. For this project, we use a refactored version of COSMO 5.0, which is able to run on hybrid GPU–CPU architectures (Fuhrer et al., 2014). The model extension was a joint effort between MeteoSwiss, the ETH-based Center for Climate Systems Modeling (C2SM), and the Swiss National Supercomputing Center (CSCS). It has been used for climate studies with 2.2 km grid spacing over Europe (Ban et al., 2014; Leutwyler et al., 2017) and is also capable of running on a near-global domain at this resolution (Fuhrer et al., 2018).

2.1.2 IFS

The Integrated Forecasting System (IFS) is the model used by the European Centre for Medium-Range Weather Forecasts (ECMWF) for its daily data assimilation and subsequent global forecasts. It is a hydrostatic model but can also be run using a nonhydrostatic extension which originally has been developed for the ARPEGE/Aladin models (Bubnová et al., 1995; Bénard et al., 2010). IFS is a spectral transform

model where temperature, wind, and surface pressure are represented in spectral space with spherical harmonics basis functions, transformed at every time step to a corresponding grid point space on a cubic-octahedral reduced Gaussian grid (Wedi, 2014; Malardel et al., 2016). Notably, all water substance variables only exist in grid point space. Semi-Lagrangian advection, physical parameterizations, and nonlinear terms are calculated in grid point space. The horizontal gradients and the Laplacian operator for horizontal wave propagation are then efficiently calculated in spectral space. The transformation between grid point space and spectral space is done by a fast Fourier transformation (FFT) in longitude and a (fast) Legendre transformation (FLT) in latitude. The spectral transforms do not scale linearly with the number of grid points and also require global communications, which means that at very high resolution the spectral transforms become a computational bottleneck of the model (Wedi et al., 2013; Schär et al., 2020). One of the reasons why the global IFS is still mostly run in hydrostatic mode is that the nonhydrostatic version uses a predictor–corrector approach that leads to more spectral transforms per time step and is therefore substantially slower. Besides the differences in the time stepping scheme, the nonhydrostatic spectral version of IFS also uses a different method for the vertical discretization then the hydrostatic IFS (finite differences instead of finite element). Recently, a nonhydrostatic core for IFS based on a finite-volume discretization (IFS-FVM) has been developed (Kühnlein et al., 2019). IFS-FVM does not require spectral transforms and achieves a higher strong scaling computational efficiency compared to the spectral model at higher resolutions. Nevertheless, the hydrostatic spectral model version of IFS is still very competitive in terms of time to solution even at the kilometer scale (e.g., Kühnlein et al., 2019; Schulthess et al., 2019; Dueben et al., 2020; Wedi et al., 2020). While it would certainly be interesting to include all three different IFS model versions into this intercomparison, the differences between them are substantial, and we consider an intercomparison among them beyond the scope of this work. A detailed comparison of the spectral hydrostatic IFS with the spectral nonhydrostatic IFS can be found in Dueben et al. (2020), who performed global simulations at 1.45 km grid spacing with these model versions in different configurations. In the present study, we have decided to focus only on the operational spectral hydrostatic model version of IFS. Therefore, when we use the term IFS in this paper, we generally refer to the operational spectral model. IFS uses an adapted version of the convection scheme by Tiedtke (1989) with improvements regarding tropical variability (Bechtold et al., 2008) and diurnal cycle (Bechtold et al., 2014). Other parameterizations include a Monte Carlo Independent Column Approximation (McICA) for radiation (Barker et al., 2008; Hogan et al., 2017) and the land surface hydrology scheme HTESSEL (Balsamo et al., 2011). IFS applies explicit diffusion to the prognostic variables in spectral space (temperature, wind,

surface pressure) with an operator that mimics spectral viscosity after Gelb and Gleeson (2001). Furthermore, some diffusion comes implicitly from the interpolation required by the semi-Lagrangian scheme in grid point space (notably for the water variables and tracers), as well as the spectral truncation due to the transformation from grid point space to spectral space, acting like a $4\Delta x$ spectral filter in the case of a cubic grid. The simulations for this project were performed with the global atmospheric model of the IFS based on Science Version 45r1. IFS documentation of the different model cycles can be found on the ECMWF website (https://www.ecmwf.int/en/publications/ifs-documentation, last access: 1 February 2021).

2.2　Numerical experiment and model setup

2.2.1　Model intercomparison

The simulations cover 2 d from 29 May 2018 00:00 UTC to 30 May 2018 with heavy thunderstorms over Europe. Both models use 1 d lead time (28 May) and are initialized with ECMWF operational analysis data at a horizontal grid spacing of $\sim$ 9 km. Since COSMO uses a different soil model (Schlemmer et al., 2018), the soil in COSMO was initialized with an average value from May/June after a 5-year spinup with COSMO 12 km by Vergara-Temprado et al. (2021). IFS is run globally, whereas COSMO is run regionally with the ECMWF operational analysis data as lateral boundary conditions on a limited area domain ranging from 361×361 ($\Delta x = 12$ km) to 1542×1542 ($\Delta x = 2.2$ km) grid points. The domain for this study is shown in Fig. 1. Simulations have been performed for a range of horizontal grid spacings for both models. The grid spacings used with COSMO are $\Delta x = 12$, 4.4, and 2.2 km, whereas the grid spacings used with IFS approximately correspond to $\Delta x = 9$, 4.5, and 2.9 km. Both COSMO and IFS usually use deep convection parameterization for coarser resolutions such as $\Delta x = 12$ km for COSMO or $\Delta x = 9$ km for IFS. However, for this study the deep convection parameterization is switched off by default and only one simulation for each model has been performed at the coarsest resolution with the deep convection parameterization on. In order to test the impact of time step in the respective simulations, each horizontal resolution setup with explicit deep convection has also been run with a smaller than usual time step. Some model properties and the respective configurations are listed in Table 1.

2.2.2　Horizontal diffusion experiment

For this experiment, COSMO has been run for the same case as above but with a varying amount of explicit diffusion from a monotonic fourth-order linear scheme with orographic limiter. This will give us some idea about the influence of diffusion on the model results and might explain some characteristic differences between IFS and COSMO. In COSMO,

Table 1. Model description and setup.

	COSMO	IFS
Numerics	Split explicit	Semi-Lagrangian and semi-implicit
Vertical velocity	Nonhydrostatic	Hydrostatic
Horizontal discretization	Rotated lat/long	Spectral and reduced Gaussian, octahedral
Resolution setups	$\Delta x = 12$ km ($0.11°$), $\Delta t \in \{90, 40\,\text{s}\}$	$\Delta x = 9$ km (TCo 1279), $\Delta t \in \{450, 240\,\text{s}\}$
	$\Delta x = 4.4$ km ($0.04°$), $\Delta t \in \{40, 20\,\text{s}\}$	$\Delta x = 4.5$ km (TCo 2559), $\Delta t \in \{240, 120\,\text{s}\}$
	$\Delta x = 2.2$ km ($0.02°$), $\Delta t \in \{20, 10\,\text{s}\}$	$\Delta x = 2.9$ km (TCo 3999), $\Delta t \in \{120, 60\,\text{s}\}$
Vertical levels	60 (up to ~ 23.5 km)[c]	137 (up to ~ 80.5 km)[c]
Convection parameterization	Deep convection off[a]	Deep convection off[b]
	Shallow convection on	Shallow convection on

[a] For $\Delta x = 12$ km, $\Delta t = 90$ s also one run with deep convection parameterization on. [b] For $\Delta x = 9$ km, $\Delta t = 450$ s also one run with deep convection parameterization on. [c] The vertical spacing is the same for all resolution setups.

fourth-order diffusion is applied on model levels by introducing an additional operator at the right-hand side of the prognostic equation, similar to the following:

$$\frac{\partial \psi}{\partial t} = S(\psi) + \texttt{diff} \cdot c_\text{d} \cdot \nabla^4 \psi \,, \qquad (1)$$

where ψ is the prognostic variable, S represents all physical and dynamical source terms for ψ, c_d is the default diffusion coefficient in the model, and $\texttt{diff}$ is the factor that can be set in order to change the strength of the computational mixing (please refer to Sect. 5.2 in Doms and Baldauf, 2018, for the exact equations including the orographic limiter). A value of $\texttt{diff} = 1$ means that the diffusion coefficient c_d remains unchanged. Any value of $\texttt{diff}$ smaller than 1 decreases the explicit diffusion coefficient, and any value larger than 1 increases explicit fourth-order diffusion strength. In our default setup for the intercomparison, COSMO has been run with no explicit fourth-order linear horizontal diffusion, which means $\texttt{diff}$ was set to 0. For this experiment, the 2.2 km setup with a time step of $\Delta t = 20$ s has been used but with numbers for $\texttt{diff}$ ranging from 0 to 4 with an increment of 0.5.

2.3 Observations

Three datasets are used for the evaluation of the model results: IMERG, RADKLIM, and IDAWEB. Comparing model results with observational data is a difficult undertaking. Next to the differences in spatial sampling (i.e., point measurement vs. grid cell averages), observations also suffer from several deficiencies (see below); therefore, different observational datasets often provide substantially different results, which is also the case in this study. Thus, observations should only be taken as a point of reference and not the absolute truth.

2.3.1 IMERG

The Integrated Multi-satellitE Retrievals for GPM (IMERG) dataset (Huffman et al., 2019b) provides worldwide half-hourly precipitation data on a $0.1° \times 0.1°$ grid by using a set of algorithms to combine satellite data and rain gauge observations into one product (Huffman et al., 2019a). IMERG incorporates satellite data from as many satellites as possible, i.e., not only the ones under the direction of the Global Precipitation Measurement (GPM) mission, in a flexible framework. The satellite data consist of passive microwave (PMW) sensors from various low-Earth-orbit platforms and infrared (IR) estimates from geosynchronous Earth-orbit satellites, as well as active radar data from the GPM satellites. The rain gauge data stem from the Global Precipitation Climate Centre (GPCC) which is operated by the German Weather Service (DWD, Deutscher Wetterdienst). The specific product that has been used by IMERG for the time period of this case study is the GPCC Monitoring Product V6 (Schneider et al., 2018). This product is based on monthly SYNOP and CLIMAT (Becker et al., 2013) data from 7000–9000 rain gauges worldwide. IMERG adjusts the accumulated monthly precipitation totals from GPCC with a gauge correction algorithm by Legates and Willmott (1990) and then calibrates the gridded multi-satellite estimate with these values. For this study, the Final version of IMERG has been used, and the half-hourly measurements were added up to hourly values in order to be consistent with the model output frequency.

2.3.2 RADKLIM

RADKLIM (Radarklimatologie) is a radar-derived and gauge-adjusted precipitation product from the German Weather Service (DWD, Deutscher Wetterdienst) that works on a 1100×900 grid over Germany with 1 km grid spacing (Winterrath et al., 2017). It uses measurements from 17 C-band weather radars (for the evaluated period only 16 radars have been in use) and approximately 2000 rain gauge stations. The method is based on the disaggregation of daily precipitation estimates from rain gauges into hourly values using radar-based estimates (Paulat et al., 2008; Wüest et al., 2010). The specific product that has been used for this study is the RW product, which uses the weighted mean of two different gauge calibration methods, from the reanalysis ver-

sion 2017.002. RADKLIM delivers hourly accumulated precipitation values for the hour from $(hh - 1):50$ to $hh:50$. This represents a slight shift compared to the model and averaged IMERG outputs, which are available for the $(hh - 1):00$ to $hh:00$ intervals. This 10 min time shift is neglected in this study. RADKLIM works very similar to RADOLAN (Bartels et al., 2004), but unlike RADOLAN it is not a real-time product. RADKLIM includes more rain gauge stations for the calibration ($\sim$ 2000 compared to $\sim$ 1300) and also possesses a more sophisticated radar artifact correction process. Therefore, RADKLIM should deliver more accurate values than RADOLAN.

Radar-based estimates of rainfall allow for a high resolution in space and time, but they are also associated with some uncertainties. Sources of errors include cluttering from other objects, attenuation, variability of the relation between reflectivity and rainfall rate (Z-R relation), beam blockage, range degradation, vertical variability of the precipitation system (i.e., bright band), and vertical air motions that increase or decrease raindrop fall speed (Villarini and Krajewski, 2010). RADKLIM uses a sophisticated radar artifact correction process which, together with the rain gauge calibration, should minimize the uncertainty due to such artifacts. Nevertheless, some artifacts might still exist, which has to be kept in mind when using the data.

An intercomparison between RADOLAN, which is very similar to RADKLIM, and IMERG can be found in Ramsauer et al. (2018), where they have found total precipitation in IMERG to be generally higher than in RADOLAN but lower in mountainous regions. This underestimation of precipitation by IMERG at higher altitudes, opposing the general overestimation in flatter terrain, was also found in studies by Wang et al. (2019), who evaluated IMERG with a dense rain gauge station network in Lhasa. They also showed that the performance of IMERG overall decreases with increasing elevation.

2.3.3 IDAWEB

IDAWEB is a web portal operated by MeteoSwiss which provides hourly precipitation measurements from roughly 1000 rain gauges from different institutions in Switzerland. Unlike IMERG or RADKLIM, it is not a gridded dataset. But due to its high density and relatively homogenous distribution of stations, it provides a good estimate of rainfall in Switzerland. The rain gauges are distributed throughout various altitude levels with the lowest one in Switzerland being at 200 m (Isole di Brissago, Brissago Islands) and the highest one at 3294 m (Piz Corvatsch). With close to 200 rain gauges at an elevation of over 1500 m, IDAWEB also provides decent coverage at high altitude in the Alpine region. IDAWEB also incorporates rain gauge measurements from a few stations outside of Switzerland, but as most of them are quite isolated, these stations were ignored for this study.

Like satellite-based or radar-based products, also rain gauge observations involve uncertainties. They suffer from various errors such as evaporation, splashing, and most importantly wind effects which usually result in a low bias. The mean undercatch for Switzerland in summer is estimated to be 7 % with exposed stations having roughly twice the bias as well-protected sites (Sevruk, 1985; Richter, 1995). A good overview of errors and error correction can be found in Sevruk (2005). While IMERG uses a gauge correction algorithm by Legates and Willmott (1990), no comparable gauge correction algorithm is applied for RADKLIM or the IDAWEB station data. However, the undercatch for heavy summer precipitation is assumed to be rather small and incorporating such possible observation errors into this analysis would be beyond the scope of this study.

3 Results

3.1 Model intercomparison

3.1.1 Precipitation pattern

We start by showing an example of the spatial precipitation distribution. Figure 1 shows accumulated hourly precipitation between 17:00 and 18:00 UTC on 29 May 2018 from different model runs and the multi-satellite product IMERG. While the location of precipitation is generally similar to the observations, there are distinct differences visible between COSMO and IFS. Most obvious is the larger amount of light precipitation in IFS compared to COSMO. Additionally, the cell structure in COSMO is more fine-grained than in IFS. These two model characteristics hold true throughout all resolutions and can also be seen in Fig. 2, which shows the precipitation patterns at the same time for the RADKLIM domain. Both runs with parameterized deep convection clearly produce less heavy precipitation than the ones with explicit deep convection. Looking at Fig. 1, one could argue that IFS 9 km (450 s) without deep convection parameterization is the closest to IMERG with regard to cluster size and location. However, this is a momentarily snapshot; while the IFS 9 km operational configuration is well-tuned, there is evidence of shortcomings and ongoing work to improve cluster size and location in the IFS model with deep convection parameterization. For the RADKLIM domain in Fig. 2, COSMO (PAR) produces much less precipitation, which is partially due to a shift in timing (precipitation falls too early in the course of the day) but also due to a generally strong underestimation of precipitation over Germany with this setup, especially for the first day (see Fig. 1 and also Table 2).

When looking at the observations only, RADKLIM and IMERG agree well on the location of the precipitation. There are, however, visible differences in intensity and spatial extent. While some of these differences might come from the different measurement and processing techniques, differ-
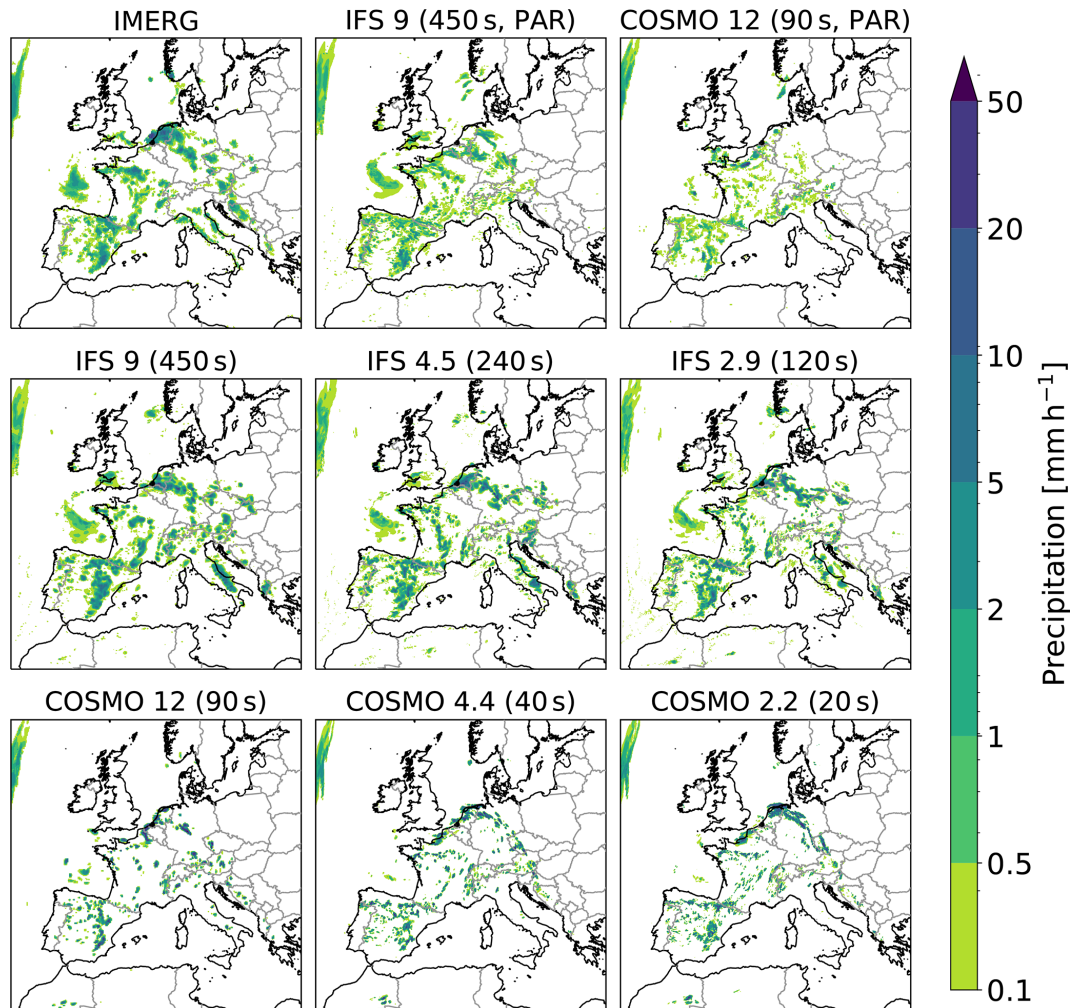
Figure 1. Accumulated hourly precipitation between 17:00 and 18:00 UTC on 29 May 2018 over the European domain. The top left panel shows precipitation data obtained from the multi-satellite product IMERG which are provided on a $0.1° \times 0.1°$ grid. The other panels show results from model runs, where the first number corresponds to the approximate grid spacing and the number in parenthesis is the respective time step. The term PAR denotes that the respective runs used parameterized deep convection.

ences will also be caused by the much higher spatial resolution of RADKLIM.

3.1.2 Precipitation intensity

The cumulative frequencies of hourly precipitation within the European domain for all 48 h are depicted in Fig. 3. Both models show an increase in frequency of precipitation ($> 0.1\,\mathrm{mm\,h^{-1}}$) but a decrease in heavy precipitation events if deep convection is parameterized. IFS produces more light precipitation than COSMO in all configurations. While the frequency of medium and heavy precipitation remain relatively constant for the COSMO runs with explicit deep convection, IFS shows increasing medium and heavy precipitation with decreasing grid spacing and time step. It can be assumed that this increase is not so much caused by the refined horizontal resolution but rather by the smaller time step.

This becomes obvious when one compares the IFS configurations with the same time step: the runs with the same time step show striking similar cumulative frequencies for different horizontal grid spacings (9 vs. 4.5 km and 4.5 vs. 2.9 km). There are three possible explanations for this strong dependence of IFS on time step:

(a) It could be that the original time step in IFS is too large to properly represent deep convective processes associated with such high vertical wind velocities (see also Sect. 3.1.5). For instance, with a 120 s time step (IFS 2.9 km) and assuming a mid-tropospheric vertical wind speed of $20\,\mathrm{m\,s^{-1}}$, an air parcel would traverse the troposphere in merely 3–4 time steps. The large time step thus implies inaccuracies, as the air parcel's trajectory and its forcing by diabatic heating cannot be fully accounted for. So while the semi-Lagrangian scheme prevents the model from developing instabilities, the large time step will likely affect the convective mesoscale
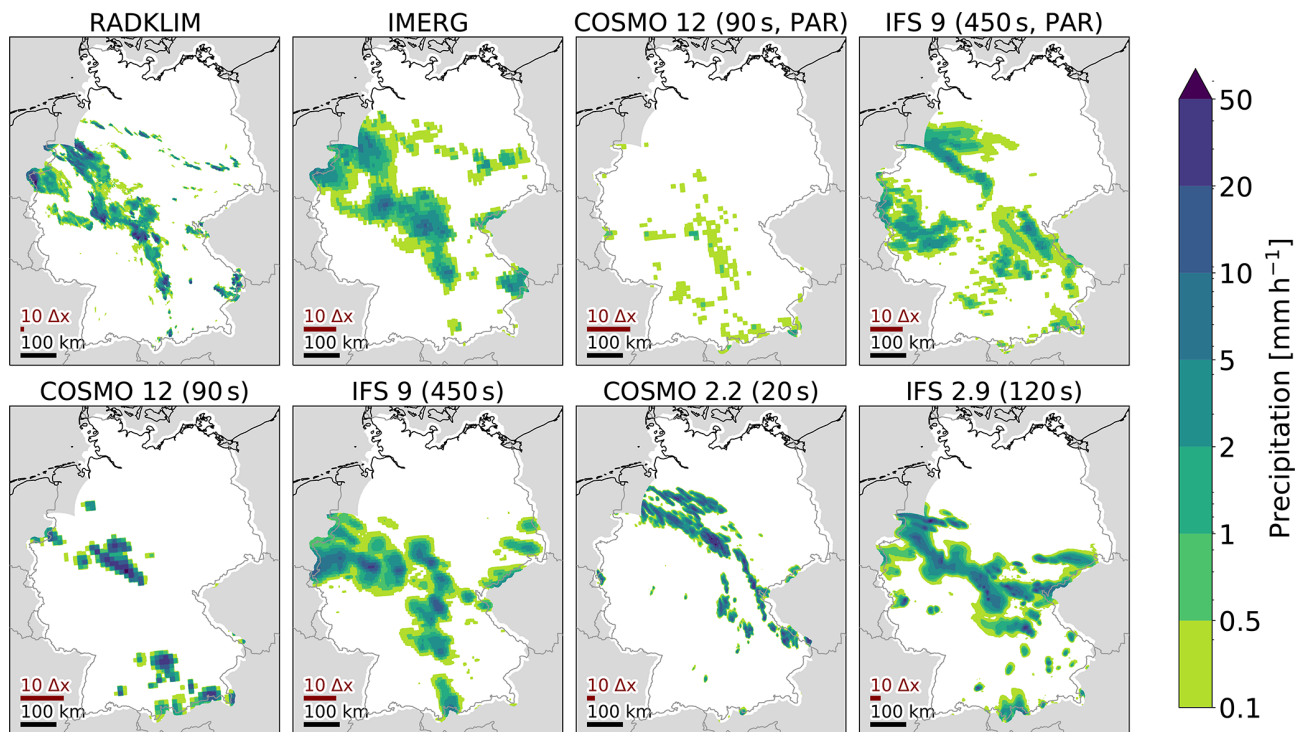
Figure 2. Accumulated hourly precipitation interpolated to the RADKLIM domain over Germany. While IMERG and the results from the model runs show precipitation from 17:00 to 18:00 UTC on 29 May 2018, the corresponding RADKLIM interval is 16:50 to 17:50 UTC (see Sect. 2.3).

dynamics, truncate extreme updrafts, and thus allow for less heavy precipitation events. As COSMO does not show such a time step sensitivity, one could argue that the time step in the COSMO simulations is already small enough and thus has no significant effect on convective mesoscale dynamics. However, also COSMO develops higher vertical velocities with a smaller time step (see Sect. 3.1.5), even though a bit less pronounced than IFS. So it is hard to imagine that this truncation of vertical updrafts is the only reason for the strong time step dependence of precipitation in IFS.

(b) By halving the time step, any subgrid-scale parameterization scheme will be called twice as often, which may affect precipitation. Notably, parameterizations such as cloud microphysics, shallow convection, or vertical mixing could experience time step sensitivity which could affect convective processes. Barrett et al. (2019) shows a strong time step sensitivity of total precipitation in an idealized setup with COSMO in combination with a two-moment microphysics scheme (see above). For our simulations, COSMO and IFS both use a single-moment scheme which shows little time step sensitivity of total precipitation but some sensitivity regarding precipitation location and magnitude in Barrett et al. (2019). While we do not see an impact in COSMO, time step sensitivity of subgrid-scale parameterization could affect IFS, where the absolute differences in time step size are generally larger.

(c) One possibility that has been investigated was the sensitivity of the interpolation error in the semi-Lagrangian scheme to time step. In semi-Lagrangian schemes, the accumulation of errors is also a function of the time step and the error of the spatial interpolation procedure (Bonaventura, 2004). Halving the time step leads to twice as many interpolations, which can potentially increase the total interpolation error, lead to more damping, and thus increase the amount of implicit diffusion. In Sect. 3.2, it is shown that more diffusion leads to more heavy precipitation events, which could thus explain the time step sensitivity in IFS. However, we do not see increased damping in IFS with a smaller time step. This becomes obvious when looking at the spectra of horizontal kinetic energy in Fig. 8, where IFS does not show stronger dissipation for the very small wavelengths with a reduced time step. The absence of such an increase in damping with smaller time step can be explained with the influence of time step on departure point in semi-Lagrangian schemes. A smaller time step will yield different departure points and thus also different damping from the quasi-cubic interpolation in IFS (see Sect. 13.6.4 in Lauritzen et al., 2011, for a more detailed discussion). We therefore consider this possibility to be less relevant when compared to (a) and (b).

Presumably, the time step dependency in IFS stems from a combination of (a), (b), and (c). But getting a better insight into the respective role of these factors would require fur-

ther studies. For example, to quantify the effect of (a), one could perform a convergence analysis with IFS by changing the time step for the dynamics while keeping the time step for the subgrid-scale parameterizations constant. Similarly, to quantify the effect of (b), the time step for the subgrid-scale parameterizations could be changed while keeping the time step for the dynamics constant. However, this is beyond the scope of this work.

Figure 4 shows the cumulative frequencies of precipitation for the RADKLIM domain and the IDAWEB stations. For the RADKLIM domain, the model output and IMERG data have been interpolated to the RADKLIM grid and for the IDAWEB stations, a nearest-neighbor interpolation has been performed. Both RADKLIM and IDAWEB show a lower frequency of precipitation ($\geq 0.1\,\mathrm{mm\,h^{-1}}$) but a higher frequency of heavy precipitation than IMERG. The higher frequency of precipitation in IMERG compared to the other observational datasets is consistent with systematic biases of the respective analysis (Ramsauer et al., 2018; Wang et al., 2019). In the RADKLIM domain, the model runs show a similar behavior as for the whole European domain. For the IDAWEB stations, it is a bit more difficult to see a clear pattern, most likely due to the small sample size in combination with the nearest-neighbor interpolation. For example, the two COSMO 2.2 km simulations show big differences in heavy precipitation. This can most probably be attributed to differences in precipitation location and subsequent different intensities obtained by the nearest-neighbor interpolation, as we do not see such a behavior for the whole European domain or the gridded RADKLIM dataset. However, both datasets, IDAWEB and RADKLIM, show that the runs with parameterized deep convection fail to produce the medium-to-heavy precipitation associated with deep convection. The cumulative frequencies of the model runs with explicit deep convection usually lie within the range of the observational values; therefore, none of these configurations seem to produce unreasonable values. In general, the distribution of the runs with higher resolutions are closer to the observations.

3.1.3 Diurnal cycle over land

Several studies have already shown that parameterized deep convection leads to a premature diurnal cycle in COSMO (Hohenegger et al., 2008; Ban et al., 2014; Leutwyler et al., 2017; Vergara-Temprado et al., 2020). Figure 5 shows that this also applies in this study. Compared to COSMO, the phase of IFS with parameterized deep convection is shifted a bit towards the later hours, which could be a result of the improvements from Bechtold et al. (2014) to the parameterization scheme by Tiedtke (1989). But still, both IFS and COSMO show a significant phase shift with parameterized deep convection. Convective precipitation lasts longer in IFS than in COSMO for all configurations. Compared to the observations from IMERG, convective precipitation in IFS lasts too long for both days, whereas in COSMO it

seems to be about right for the first day but too short for the second day. A very interesting aspect is the dependence of the diurnal cycle on spatial and temporal resolution. All runs, except the ones that already use a rather small time step (COSMO 4.4 and 2.2), show an earlier development and also decay of convective precipitation when reducing the time step. There are some signs of convergence for COSMO at 4.4 km grid spacing and also for IFS, where the 4.5 km run with $\Delta t = 120\,\mathrm{s}$ and the two 2.9 km runs are quite close. This convergence of the diurnal cycle around 4 km is somewhat consistent with findings from Langhans et al. (2012) and Panosetti et al. (2019), who simulated 9 convective summer days with COSMO over the Alps and Germany with grid spacings ranging from 8.8 km to 550 m. In their studies, the peak of the diurnal cycle then shifts back, again more towards the evening, with 1.1 km and 550 m. It is possible that, by increasing the resolution even further, a similar effect could be seen in this case. The shift of the diurnal cycle with higher temporal resolution for the explicit COSMO 12 km is also consistent with the results from Panosetti et al. (2019), who found a shift of the 8.8 km run but none for the 4.4 km run by using $\Delta t = 5\,\mathrm{s}$ instead of the original time step of 80 and 40 s, respectively. It is not entirely clear what causes this dependence of the diurnal cycle on temporal resolution. It could be that a large time step leads to some truncation in the dynamics and thus to a delay in inhibition. But it is also likely that this is related to the changed frequency of application of the subgrid-scale parameterizations. A detailed analysis of the causes for this behavior would certainly be an interesting subject for further studies.

For COSMO, Baldauf et al. (2011) has shown a high sensitivity of convection initiation to the Blackadar length scale. But as the Blackadar length scale has been kept constant for all COSMO configurations ($l_\infty = 150\,\mathrm{m}$) and as the same effect regarding resolution can be seen in IFS, which uses a different parameterization for turbulent transport (ECMWF, 2018), it can be ruled out as a factor in this case.

3.1.4 Total precipitation

Total precipitation during the 2 d has been analyzed for four domains: the whole European domain, the land part of the European domain, the RADKLIM domain, and the IDAWEB stations. The results are summarized in Table 2.

For the whole European domain, all COSMO runs show clearly less precipitation than IMERG. And while all COSMO simulations with explicit deep convection produce about the same amount of precipitation, the one with parameterized deep convection is clearly an outlier with even less precipitation. The IFS runs with explicit deep convection show about the same amount of precipitation as IMERG. Also here, the run with parameterized deep convection shows significantly less precipitation than the explicit ones, similar to the results of the global simulations with IFS by Dueben et al. (2020). This is most probably due to the miss-
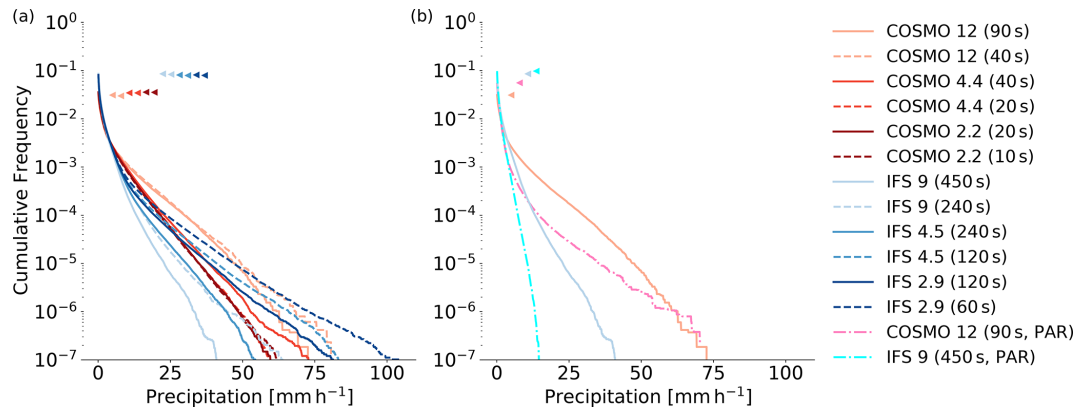
Figure 3. Cumulative frequency of accumulated hourly precipitation over the European domain for the whole 48 h period. Panel **(a)** shows all runs with explicit deep convection and two different time steps per spatial resolution. Panel **(b)** shows the coarser runs with parameterized (PAR) and with explicit deep convection. The triangles indicate the frequency of precipitation $\geq 0.1\,\mathrm{mm\,h^{-1}}$ of the respective simulations.
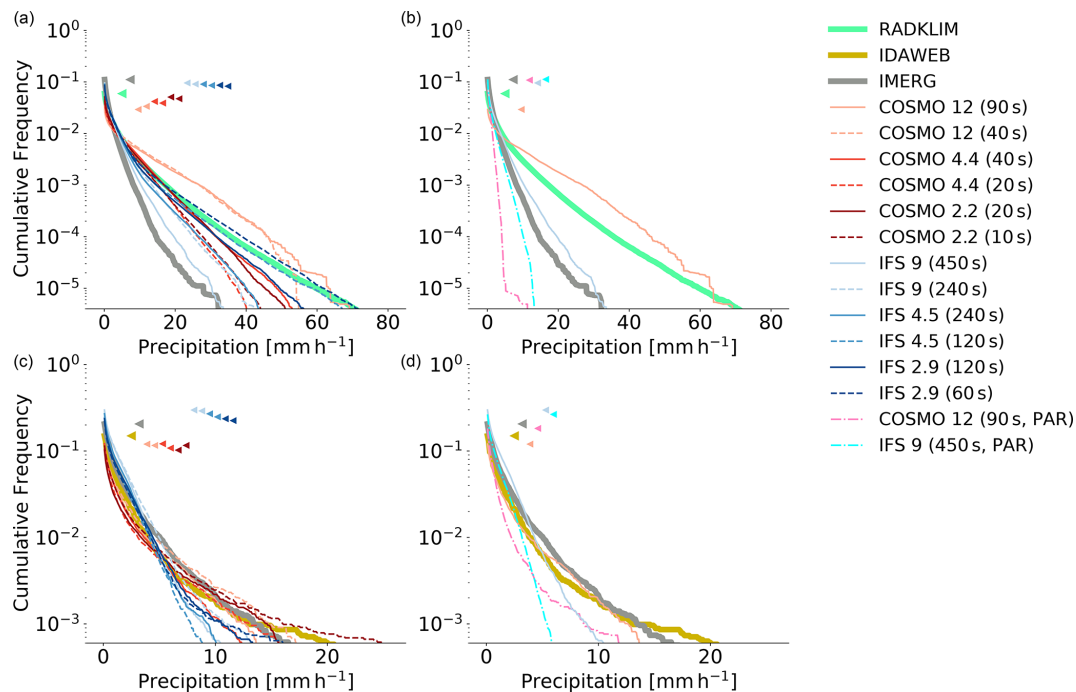

Figure 4. Same as Fig. 3 but for the RADKLIM domain **(a, b)** and IDAWEB station data **(c, d)**. Note the different scales for the two domains.

ing medium-to-heavy precipitation in the parameterized runs as shown in Sect. 3.1.2.

If one looks only at the precipitation over land, COSMO is much closer to IMERG, while the values from IFS are clearly larger. The effect of parameterized deep convection for both models is the same as for the whole European domain but even more distinct as the larger part of deep convection is happening over land. Moreover, IFS shows a clear sensitivity to time step with the amount of precipitation increasing with decreasing time step.

One of the properties of hydrostatic systems is supposed to be the overestimation of convective precipitation amount and

area compared to nonhydrostatic systems (Kato and Saito, 1995; Kato, 1997). When looking at total precipitation over land and comparing it with values from IMERG, it looks like IFS is overestimating convective precipitation. Additionally, the overestimation seems to get worse with increasing resolution which is consistent with findings by Kato (1997). However, it is not clear whether this effect can be purely attributed to the hydrostatic core, as there are other factors, notably the subgrid-scale parameterizations, to consider.

Total precipitation in the RADKLIM domain and at the IDAWEB stations has to be interpreted cautiously as both domains are rather small and the simulations cover only 48 h.
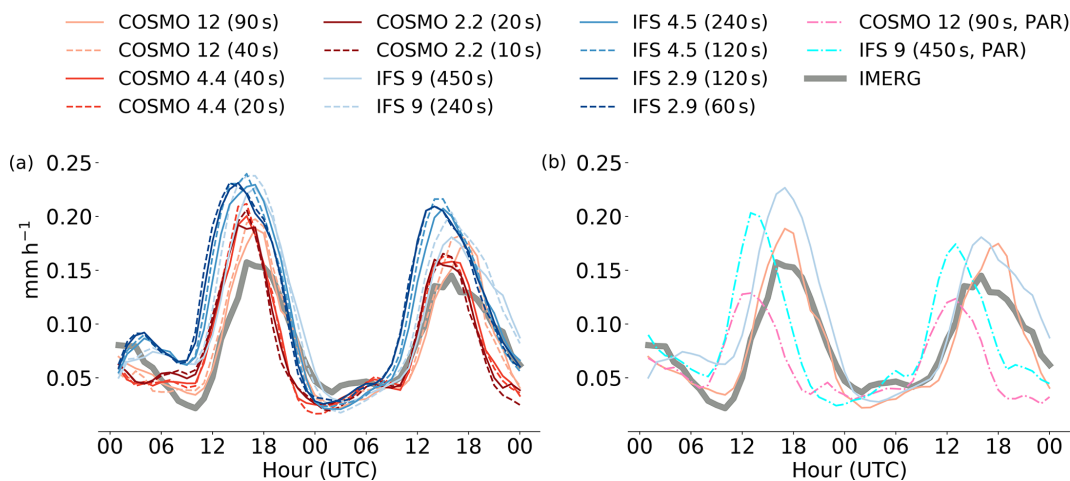
Figure 5. Diurnal cycle of precipitation over land in the European domain from 29 May 2018 01:00 UTC to 31 May 2018 00:00 UTC. The plot shows accumulated hourly precipitation, meaning that at 29 May 2018 01:00 UTC it shows the precipitation accumulated from 29 May 2018 00:00 UTC to 29 May 2018 01:00 UTC. Panel **(a)** shows all runs with explicit deep convection and two different time steps per spatial resolution. The results from the runs with the shorter time step are represented by the dashed lines. Panel **(b)** shows the coarser runs with parameterized (PAR) and with explicit deep convection. Both panels also show the values from the Integrated Multi-satellitE Retrievals for GPM (IMERG) dataset for our domain.

Table 2. Mean precipitation per 48 h in the different domains. The bold numbers represent the reference values of the respective domain that have been used to calculate the differences in percent.

	European domain[a]	European domain (land)[a]	RADKLIM domain[b]	IDAWEB stations[c]
IMERG	**3.84 mm (48 h)$^{-1}$**	**3.86 mm (48 h)$^{-1}$**	−6.2 %	+37.0 %
RADKLIM	–	–	**6.93 mm (48 h)$^{-1}$**	–
IDAWEB	–	–	–	**9.58 mm (48 h)$^{-1}$**
COSMO 12 (90 s, PAR)	−37.6 %	−22.0 %	−53.1 %	−29.5 %
COSMO 12 (90 s)	−26.3 %	+0.0 %	−0.8 %	−16.0 %
COSMO 12 (40 s)	−26.6 %	+2.8 %	+9.4 %	−4.4 %
COSMO 4.4 (40 s)	−26.9 %	−2.4 %	−11.6 %	−22.6 %
COSMO 4.4 (20 s)	−28.2 %	−2.0 %	−24.3 %	−30.0 %
COSMO 2.2 (20 s)	−27.5 %	−1.8 %	−6.9 %	−29.2 %
COSMO 2.2 (10 s)	−29.2 %	−3.7 %	−17.8 %	−9.8 %
IFS 9 (450 s, PAR)	−17.1 %	+5.6 %	−29.2 %	+17.6 %
IFS 9 (450 s)	−4.9 %	+28.0 %	−2.1 %	+73.4 %
IFS 9 (240 s)	−0.9 %	+35.4 %	+16.4 %	+86.2 %
IFS 4.5 (240 s)	−1.1 %	+34.5 %	+7.4 %	+50.9 %
IFS 4.5 (120 s)	−0.6 %	+37.3 %	+11.9 %	+31.2 %
IFS 2.9 (120 s)	+0.3 %	+40.1 %	+14.3 %	+37.2 %
IFS 2.9 (60 s)	+0.1 %	+40.4 %	+13.4 %	+28.7 %

[a] Interpolated on COSMO 2.2 km grid. [b] Interpolated on RADKLIM 1 km grid. [c] Nearest-neighbor interpolation to IDAWEB stations.

But the numbers support the findings from the European domain in the sense that IFS seems to overestimate precipitation while COSMO generally underestimates it. Also, the precipitation-reducing effect of parameterized deep convection is visible for both domains.

3.1.5 Cumulative frequency of vertical winds

Figure 6 shows the cumulative frequencies of vertical wind velocity at 850 and 500 hPa. The frequencies were calculated from the instantaneous vertical velocities at every full hour during the 48 h simulation period. While the distribution on the 850 hPa level is almost symmetric in COSMO and for the coarser IFS runs, the updrafts at the 500 hPa become much stronger than the downdrafts. This property is

consistent with a cross section of a multicell thunderstorm produced by COSMO in Fig. 7. At 850 hPa, the updrafts are not really that strong due to the proximity to the planetary boundary layer. At 500 hPa, we are well above the level of free convection and the updraft velocities become very high in this area. In general, downdrafts are more frequent than updrafts in both models and on both levels, but they never develop the strength of the deep convective updrafts. Especially in IFS, the values for the downdrafts are quite low.

The profound impact of deep convection parameterization on the vertical motions in the atmosphere can be seen on the two panels on the left-hand side of Fig. 6. Both the updraft and downdraft velocities are much smaller with parameterized deep convection. This effect is most pronounced for the IFS updrafts at 500 hPa where the parameterization leads to very low updrafts which is also consistent with the lack of heavy precipitation in this configuration (see Fig. 3).

Both models show some sensitivity to horizontal resolution and the updraft velocities at 500 hPa are also comparable between the respective horizontal resolutions. The time step sensitivity seems to be more pronounced in IFS at both levels, which we interpret as resulting from the larger vertical motion in combination with a large time step. Nevertheless, the updraft velocities for the 4.5 and 2.9 km runs of the hydrostatic IFS are similar to those of the nonhydrostatic COSMO runs with 4.4 and 2.2 km grid spacing, respectively. Hence, the presumption that the vertical velocities could become unrealistically high due to the violation of the hydrostatic assumption at these resolutions cannot be confirmed. It is not clear to what extent the large time step of IFS influences these results, but results from Dueben et al. (2020) even show a reduction of updraft velocities with the hydrostatic version of IFS at 1.45 km horizontal grid spacing, when reducing the time step from 120 to 30 s.

Another interesting aspect is the disparity in downdraft velocities between IFS and COSMO. The downdraft velocities in IFS are significantly lower for the explicit runs compared to the corresponding COSMO runs. At the same time the probability of having a downdraft is higher in IFS than in COSMO. The lower downdraft velocities in IFS could be related to the hydrostatic formulation of the governing equations, as the results from Dueben et al. (2020) generally show higher downdraft values when switching to the nonhydrostatic implementation of IFS. But also diffusion seems to play a critical role in downdraft velocities and might explain the differences between the two models, as shown in Sect. 3.2.

3.1.6 Energy spectra

While kinetic energy spectra are generally not used as a measure of a model's skill, they can be useful in order to determine whether a model is able to reproduce the observed dynamics of the atmosphere (Skamarock, 2004). Observational analysis from Nastrom and Gage (1985) showed a transi-

tion of the kinetic energy spectra from a k^{-3} dependence at the large scale, characteristic of two-dimensional turbulence, to a $k^{-5/3}$ dependence at the mesoscale, with k being the wavenumber. These results have been confirmed by other studies such as Lindborg (1999), Cho et al. (1999a), and Cho et al. (1999b). The two upper panels of Fig. 8 show the power spectral density (PSD) of horizontal kinetic energy E_{kin} for all model runs. With deep convection parameterization off, both models are able to nicely reproduce the expected slopes. The amplitude and the location of the transition zone at around 400 km wavelength are almost identical for all explicit simulations and also agree very well with results from Skamarock and Klemp (2008), who performed numerical studies using the Advanced Research WRF (ARW) model. Deep convection parameterization is clearly influencing the dynamics in the sense that the $k^{-5/3}$ dependence in the mesoscales cannot be reproduced. For horizontal kinetic energy at 500 hPa, both models show very little dependence on time step, similar to the results from Dueben et al. (2020). COSMO seems to conserve a bit more kinetic energy at smaller wavelengths while IFS shows stronger dissipation at these scales. Malardel and Wedi (2016) examined kinetic energy spectra produced by IFS and identified subgrid-scale parameterizations (notably surface drag and momentum vertical mixing) as the major contributors to dissipation in IFS. They have also found that differences in orographic filtering will affect the energy transfer. It is likely that the differences between COSMO and IFS in dissipation rate for smaller wavelengths are caused by a combination of different factors.

So while time step seems to have little influence on the horizontal kinetic energy spectra, it certainly has an influence on the vertical wind spectra, as the lower right panel of Fig. 8 shows. For most runs a reduction in time step leads to significantly more energy throughout all wavelengths. This is most pronounced for the pairs of simulations with larger time steps, but even for the runs with smaller time steps it leads to an increase in power, especially for the smaller wavelengths. Similar to the horizontal kinetic energy spectra, COSMO conserves more energy in the smaller wavelengths than IFS. The effect of parameterizing deep convection seems to be even more drastic for the vertical than for the horizontal winds: the amplitude is clearly reduced throughout all wavelengths for both models.

The amplitude and shape of the power spectral densities of w from the explicit runs seem to mostly agree with values from other observational and numerical studies (Bacmeister et al., 1996; Gao and Meriwether, 1998; Callies et al., 2016; Schumann, 2019; Panosetti et al., 2019). The spectra of the configurations used in this study all follow a slope of roughly $k^{-1/5}$. This seems to be a realistic value (Liu, 2019), even though there is quite a bit of variability in the aforementioned studies, indicating that probably there is some dependence on the specific weather situation, altitude, and regional climate considered.
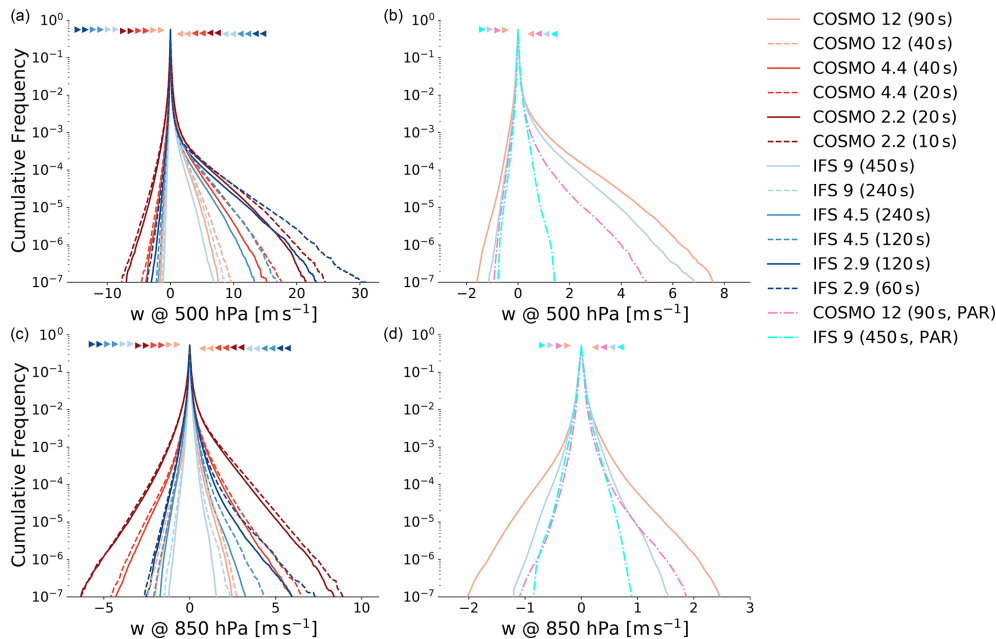
Figure 6. Cumulative frequency of vertical wind velocity on the 500 hPa (upper row) and 850 hPa pressure levels (lower row) with the same layout as in Fig. 3. The triangles indicate the height of the respective starting points, that is, the frequency of downdrafts ($< 0\,\mathrm{m\,s^{-1}}$) and updrafts ($> 0\,\mathrm{m\,s^{-1}}$). Note the different scales on the horizontal axes.
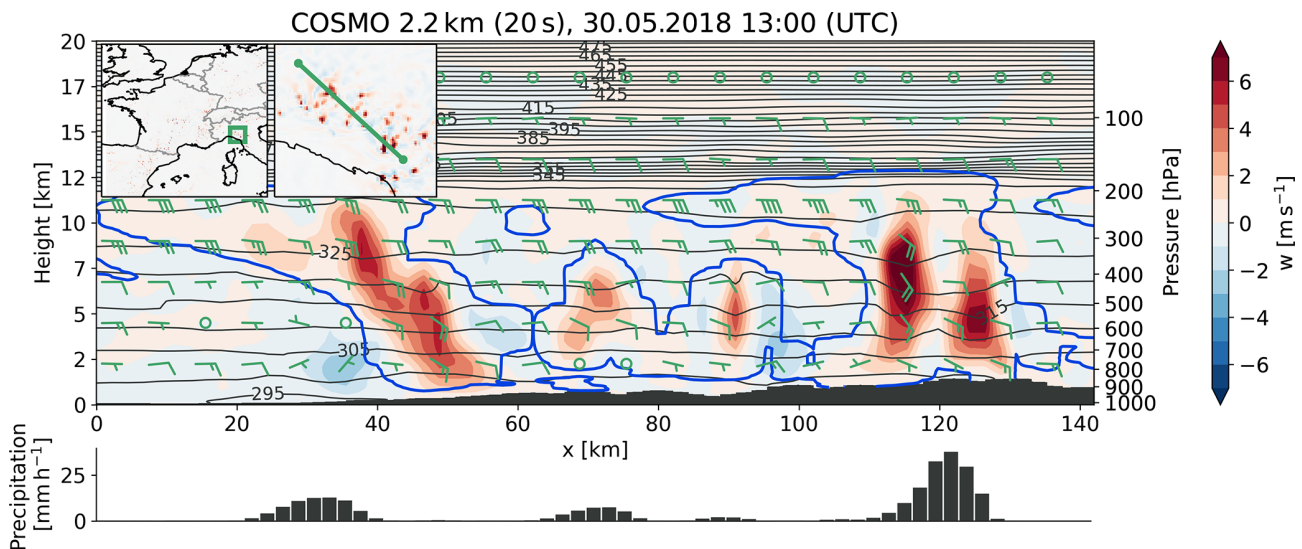

Figure 7. Cross section of convective cells over Northern Italy produced by COSMO 2.2 km. The shading represents the vertical wind speed, the green barbs the overall wind direction and velocity, and the blue contour the clouds (cloud water + cloud ice content $> 0.01\,\mathrm{g\,kg^{-1}}$). The black contours represent potential temperature, and the bar chart at the bottom shows the location and intensity of accumulated precipitation over the last hour.

3.2 The effect of horizontal diffusion

Figure 9 shows the effect of horizontal diffusion on precipitation and vertical velocity in COSMO. Rather counterintuitively, additional explicit diffusion leads to a decrease of light precipitation and a significant increase in heavy precipitation. This increase in heavy precipitation events could ex-

plain some of the difference between the COSMO 2.2 km and IFS 2.9 km simulations shown in Fig. 3, where IFS with $\Delta t = 60\,\mathrm{s}$ also reaches values of $100\,\mathrm{mm\,h^{-1}}$ and more.

The impact of horizontal diffusion on vertical wind velocities is not as prominent as for precipitation, but nevertheless there is a clear pattern. Both updraft and downdraft velocities are reduced with more horizontal diffusion. The
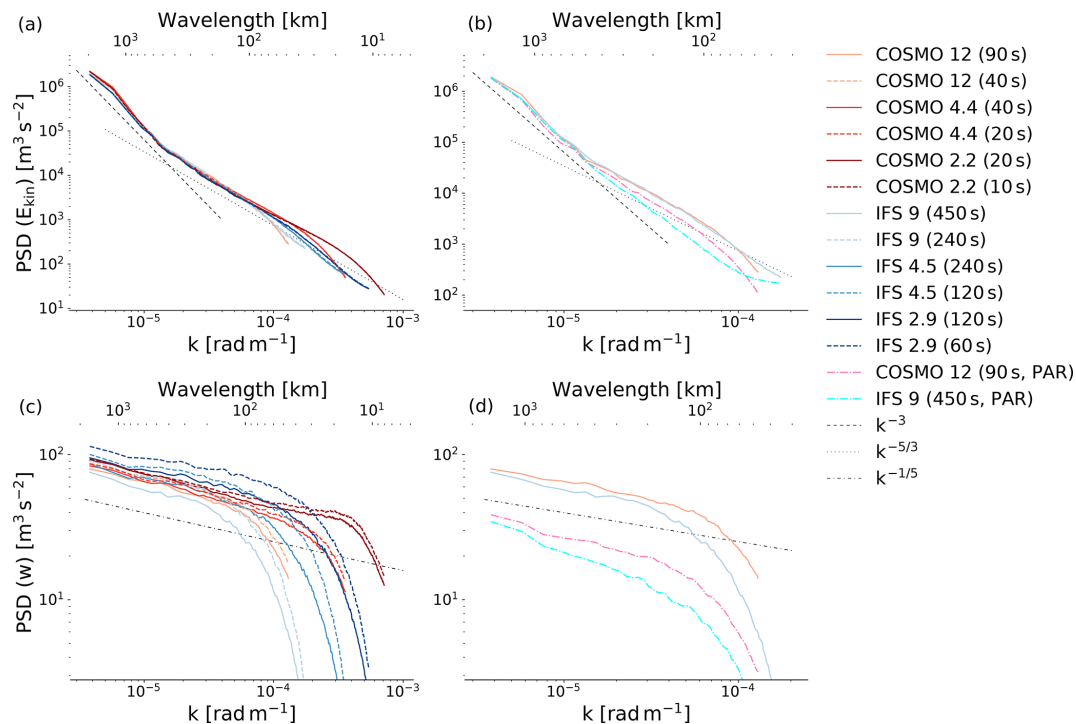
Figure 8. Power spectral density (PSD) plots for horizontal kinetic energy and w at 500 hPa (**a–d**). Panels (**a**) and (**c**) show PSDs of all runs with explicit deep convection. Panels (**b**) and (**d**) show PSDs for the coarser runs with parameterized (PAR) and explicit deep convection. Data points for wavelength smaller than $4\Delta x$ or larger than $L/2$, where L is the domain width, have been cut off. Note the different scales on the vertical and horizontal axes.

relative change is most noticeable for the downdrafts, again with the diffusive configurations showing some similarities to the behavior of IFS in Fig. 6. So increased horizontal diffusion could at least be a possible cause of the, compared to COSMO, low downdraft velocities in IFS at the 500 hPa level. The additional diffusion does not really change the downdraft velocities on the 850 hPa level (not shown) but reduces the updraft velocities and brings them closer to the corresponding profiles from IFS.

Figure 10 shows a snapshot of vertical wind velocities at 500 hPa and precipitation from a multicell thunderstorm over the area around the Netherlands. Without explicit horizontal diffusion, the cells in COSMO are much smaller than the ones in IFS. But with a large amount of explicit horizontal diffusion, the cells look quite similar to the ones in IFS in terms of size and shape. The relatively large horizontal extent of convective cells in IFS could be a reason why the hydrostatic approximation still seems to work quite well even at a horizontal grid spacing of 2.9 km. So one could argue that the rather diffusive properties of the model prevent it from entering into a nonhydrostatic regime, where the vertical extent of buoyant cells becomes larger than their horizontal extent.

This observed increase of convective cell size and heavy precipitation with additional horizontal diffusion is very similar to the results in Ricard et al. (2013) with AROME. Malardel and Ricard (2015) increased diffusion in the con-

vergent parts of the flow and reduced it in the divergent parts in order to improve the conservation property of the scheme used in IFS, AROME, and HARMONIE, which lead to a decrease of heavy precipitation. Unlike in Malardel and Ricard (2015), the additional diffusion is applied everywhere in our experiment. It is not clear how the results would change with COSMO, if for example diffusion would be applied only to the convergent or divergent part of the flow, but answering that is beyond the scope of this study and would require further investigations.

One of the most important conclusions from the COSMO diffusion experiments is the evidence that horizontal diffusion in the governing equations does not act to simply smooth the precipitation field (which would weaken and broaden all cells, but not significantly change their number). Rather it appears that diffusion more fundamentally affects the dynamics: with higher diffusion, the available convective available potential energy (CAPE) is consumed by substantially fewer but broader updrafts (Fig. 10). In terms of peak vertical velocity, however, the cells weaken, and one wonders why the peak hourly precipitation rates increase so strongly (Fig. 9). We think that the increase of peak precipitation is owed to an accumulation effect. As the cells are much broader, the precipitation footprint at the surface will take longer while propagating over an affected grid point. Evidence for such
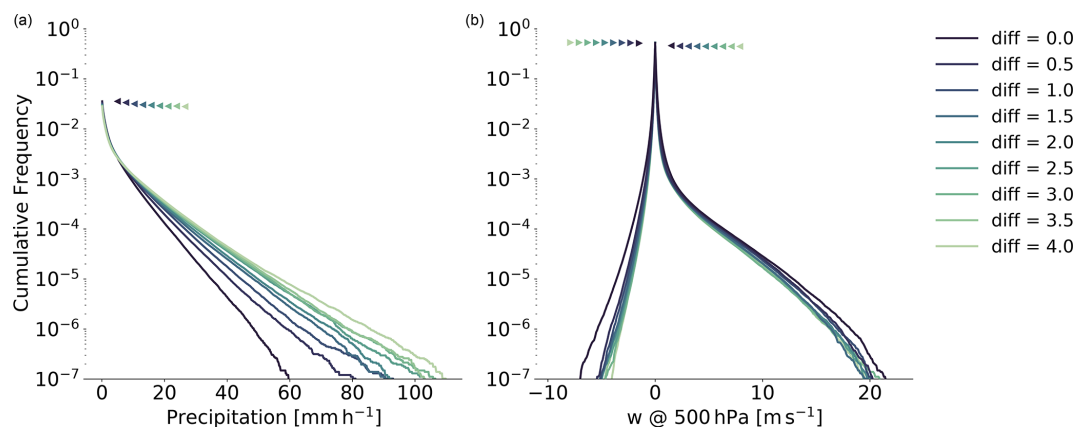
Figure 9. Cumulative frequency of hourly precipitation **(a)** and vertical wind speed at 500 hPa **(b)** for runs with COSMO 2.2 km and different values for explicit horizontal diffusion of wind, temperature, pressure, and moisture variables.

accumulation effects can be seen in the elongated hourly precipitation signatures in Fig. 10.

Figure 11 shows the power spectral densities of the COSMO diffusion experiments, and it is obvious that more diffusion leads to more dampening near the short-wave cutoff. In fact, the spectra from COSMO with substantial explicit diffusion look quite similar to the ones obtained from IFS with 2.9 km grid spacing in Fig. 8.

4 Conclusions

IFS produces more light precipitation than COSMO in all configurations and generally produces more precipitation. For both models, parameterized deep convection leads to more light precipitation but less medium-to-heavy precipitation. With explicit deep convection, the cumulative frequencies in COSMO are quite constant with regard to horizontal resolution and time step. This is not the case for IFS, which shows an increasing amount of heavy precipitation with increasing resolution. However, the deciding factor for the precipitation frequencies in IFS seems to be the time step. IFS runs with a smaller time step all lead to significantly more heavy precipitation than the respective runs with larger time step. It is not entirely clear how much this behavior is an effect of time step on resolved dynamics or the subgrid-scale parameterizations and their coupling. It is possible that a combination of these factors contribute to this time step sensitivity of precipitation intensities.

The comparison of model results with the three observational datasets IMERG, RADKLIM, and IDAWEB showed that both model's runs with explicit deep convection seem to be in the range of realistic values when it comes to precipitation intensities. In contrast, both runs with parameterized deep convection failed to reproduce the medium-to-heavy precipitation that could be observed during these 2 d and thus also produced significantly less precipitation.

Resolution and time step size also have an effect on the diurnal cycle of precipitation over land. Higher spatial and temporal resolutions seem to lead to an earlier onset and peak of precipitation. While we see a convergence of the diurnal cycle already at 4.4 km grid spacing in COSMO, IFS only shows signs of convergence at the highest resolution with 2.9 km grid spacing, most probably still due to the relatively large time step sizes of 120 and 60 s. Furthermore, this study also reinforces the evidence that parameterized deep convection leads to a much earlier onset and peak in the diurnal cycle. However, besides the two coarsest runs (COSMO 12 km and IFS 9 km) with explicit deep convection, all runs seem to have a too early phase in the diurnal cycle when compared with observations from the multi-satellite product IMERG.

The redistribution of heat and moisture due to parameterized deep convection has a distinct effect on the vertical velocities, leading to lower values for the downdrafts and especially the updrafts. From the runs with explicit deep convection, the respective updraft values at the 500 hPa level were quite similar between the nonhydrostatic COSMO and the hydrostatic IFS. This indicates that the hydrostatic approximation at a grid spacing of around 2–3 km still works well and does not lead to too high updraft values. However, the downdraft values in IFS are significantly lower than in COSMO throughout almost all simulations. This could be a characteristic of hydrostatic models (see for example Dueben et al., 2020) or also be caused by enhanced diffusion in IFS compared to COSMO. However, this would require further investigation.

The influence of time step on wind velocities does not seem to be very crucial for the horizontal winds and both models show almost no change in the spectra of horizontal kinetic energy with different time steps. The vertical winds, however, are clearly influenced by the time step. This is visible in changes in spectra and also frequency distributions where a large time step seems to suppress high vertical velocities. The importance of resolving all these high velocities,
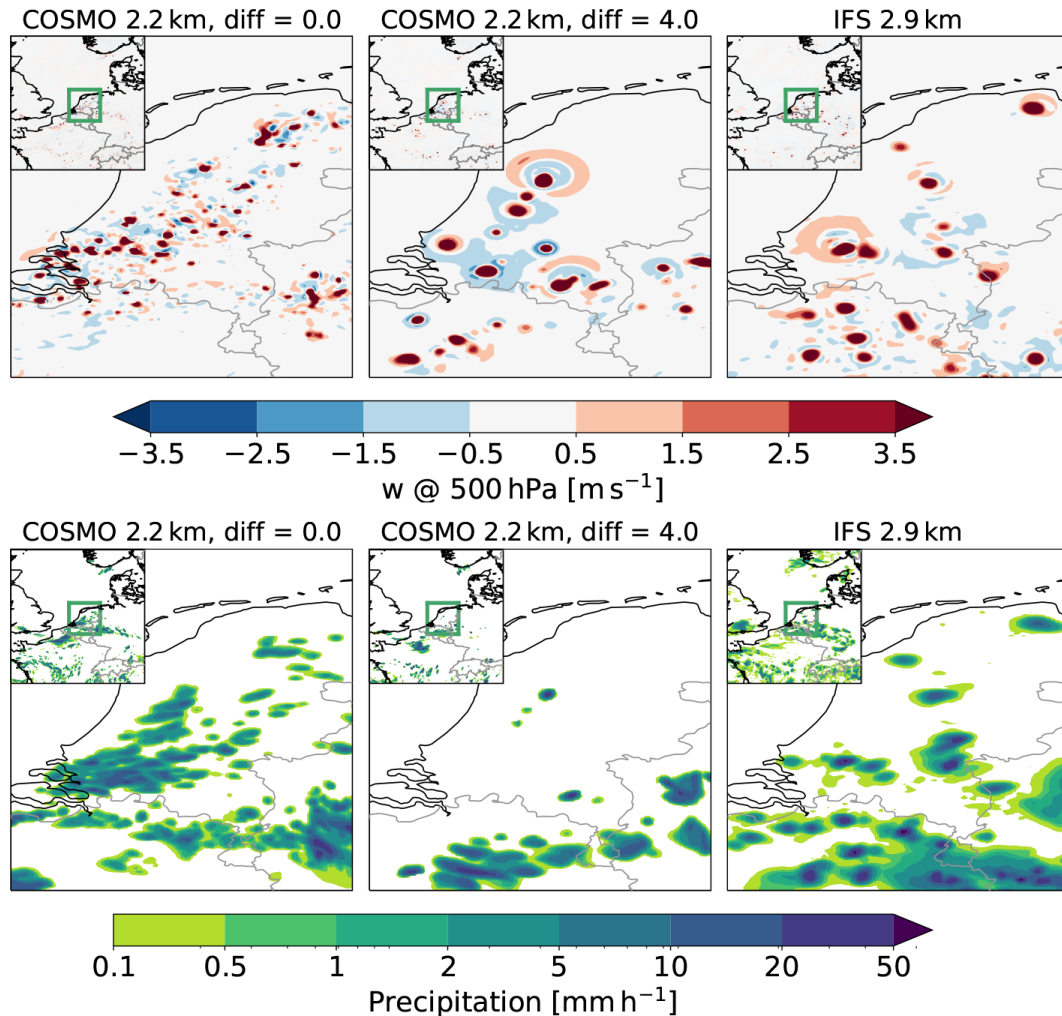
Figure 10. Vertical wind at 500 hPa (upper row) and accumulated hourly precipitation (lower row) over the Netherlands on 29 May 2018 at 14:00 UTC. The first column shows the values from COSMO 2.2 km without explicit horizontal diffusion, while the middle column shows the results from a simulation with additional diffusion. The right column shows to the values obtained from the IFS 2.9 km simulation with $\Delta t = 120\,\mathrm{s}$.
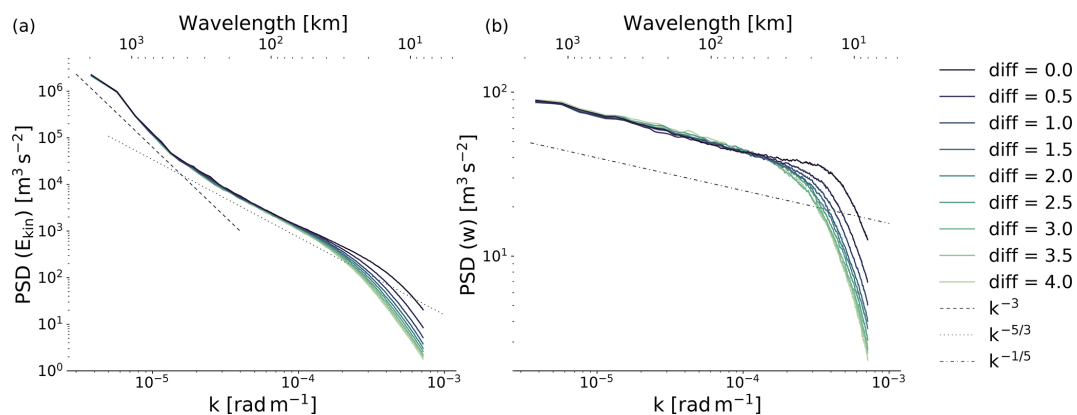

Figure 11. Power spectral density (PSD) plots for horizontal kinetic energy and w at 500 hPa for runs with COSMO 2.2 km and different values for explicit horizontal diffusion of wind, temperature, pressure, and moisture variables. Note the different scales on the vertical axes.

compared to the significant additional computational costs involved with a smaller time step, is up for debate and probably also depends on application and purpose of the simulation.

Increasing horizontal diffusion in COSMO leads to more medium and heavy precipitation, making the precipitation frequency profile in this range look similar to the ones from IFS. Furthermore, more horizontal diffusion also leads to a reduction of downdraft velocities at the 500 hPa level and thus also makes the vertical velocity profiles of COSMO and IFS more akin. The added diffusion generally leads to fewer convective cells while increasing the horizontal extent of these cells. This could be a reason why the hydrostatic approximation still seems to work quite well even at a grid spacing of around 2–3 km, as the relatively large horizontal width of the cells might prevent them from entering the non-hydrostatic regime where the vertical extent of the buoyant cells becomes larger than the horizontal extent. But while this sensitivity to dissipation certainly would need a more detailed investigation, it seems to explain some of the characteristic differences between COSMO and IFS.

Given the significant structural differences between the two models, it is very difficult to confidently attribute differences in the shown results to specific model properties. While this study is able to give some indications, it also stimulates further research. For example, is the sensitivity in heavy precipitation of IFS with regard to time step size mainly a dynamical effect (with the normal time step being too large to properly resolve the updrafts) or rather an effect from the increased calling frequency of the subgrid-scale parameterization schemes? It would be intriguing to only vary the time step of the dynamics while leaving the time step for the physical parameterizations constant (or the other way around) in order to be able to answer this question. A hypothesis in this study is that we see an increase in heavy precipitation with more horizontal diffusion mainly due to an accumulation effect of the larger convective cells and not necessarily due to heavier instantaneous precipitation. This could be verified by a similar test configuration as for our diffusion experiment but with a focus on instantaneous precipitation rates with a high output frequency. Regarding the validity of the hydrostatic approximation, the current study supports a view that it is still suitable for a grid spacing of around 2–3 km. But is this mainly due to the rather diffusive behavior of IFS with its relatively large convective cells? A study with more focus on convective cell size rather than grid spacing could probably answer this (for example similar to Miyamoto et al., 2013; Jeevanjee, 2017). What is the role of differences in time step, numerical methods, or subgrid-scale parameterizations between the hydrostatic IFS and the nonhydrostatic COSMO when it comes to updraft velocities? Investigating this question in more detail would require a study with two models that have nearly identical numerics and use the same physical parameterizations. For IFS, a prerequisite for such a comparison would be to try bringing the nonhydrostatic spectral IFS closer to the operational hydrostatic IFS with regard to the time stepping scheme, the vertical discretization, and the physics–dynamics coupling.

Author contributions. CZ, NPW, and CS designed the experiments. CZ performed the COSMO model simulations and NPW the corresponding IFS simulations. CZ performed the analysis of model output and observations with supervision from CS and NB, as well as technical support from PDD and NPW. NPW, PDD, and CS were strongly involved in the discussion of the results. CZ wrote the paper with input from all other co-authors.

Competing interests. The authors declare that they have no conflict of interest.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Bacmeister, J. T., Eckermann, S. D., Newman, P. A., Lait, L., Chan, R. K., Loewenstein, M., Proffitt, M. H., and Gary, B. L.: Stratospheric horizontal wavenumber spectra of winds, potential temperature, and atmospheric tracers observed by high-altitude aircraft, J. Geophys. Res.-Atmos., 101, 9441–9470, https://doi.org/10.1029/95JD03835, 1996.

Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M., and Reinhardt, T.: Operational Convective-Scale Numerical Weather Prediction with the COSMO Model: Description and Sensitivities, Mon. Weather Rev., 139, 3887–3905, https://doi.org/10.1175/MWR-D-10-05013.1, 2011.

Balsamo, G., Pappenberger, F., Dutra, E., Viterbo, P., and van den Hurk, B.: A revised land hydrology in the ECMWF model: a step towards daily water flux prediction in a fully-closed water cycle, Hydrol. Process., 25, 1046–1054, https://doi.org/10.1002/hyp.7808, 2011.

Ban, N., Schmidli, J., and Schär, C.: Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations, J. Geophys. Res.-Atmos., 119, 7889–7907, https://doi.org/10.1002/2014JD021478, 2014.

Barker, H. W., Cole, J. N. S., Morcrette, J.-J., Pincus, R., Räisänen, P., von Salzen, K., and Vaillancourt, P. A.: The Monte Carlo Independent Column Approximation: an assessment using several global atmospheric models, Q. J. Roy. Meteor. Soc., 134, 1463–1478, https://doi.org/10.1002/qj.303, 2008.

Barrett, A. I., Wellmann, C., Seifert, A., Hoose, C., Vogel, B., and Kunz, M.: One Step at a Time: How Model Time Step Significantly Affects Convection-Permitting Simulations, J. Adv. Model. Earth Sy., 11, 641–658, https://doi.org/10.1029/2018MS001418, 2019.

Bartels, H., Weigl, E., Reich, T., Lang, P., Wagner, A., Kohler, O., Gerlach, N., and MeteoSolutions GmbH: Projekt RADOLAN – Routineverfahren zur Online-Aneichung der Radarniederschlagsdaten mit Hilfe von automatischen Bodenniederschlagsstationen (Ombrometer), Tech. rep., Deutscher Wetterdienst, Hydrometeorologie, available at: https://www.dwd.de/DE/leistungen/radolan/radolan_info/abschlussbericht_pdf.pdf?__blob=publicationFile&v=2 (last access: 13 July 2021), 2004.

Bechtold, P., Köhler, M., Jung, T., Doblas-Reyes, F., Leutbecher, M., Rodwell, M. J., Vitart, F., and Balsamo, G.: Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales, Q. J. Roy. Meteor. Soc., 134, 1337–1351, https://doi.org/10.1002/qj.289, 2008.

Bechtold, P., Semane, N., Lopez, P., Chaboureau, J. P., Beljaars, A., and Bormann, N.: Representing equilibrium and nonequilibrium convection in large-scale models, J. Atmos. Sci., 71, 734–753, https://doi.org/10.1175/JAS-D-13-0163.1, 2014.

Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., and Ziese, M.: A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present, Earth Syst. Sci. Data, 5, 71–99, https://doi.org/10.5194/essd-5-71-2013, 2013.

Bénard, P., Vivoda, J., Mascaronek, J., Smolíková, P., Yessad, K., Smith, C., Brožková, R., and Geleyn, J. F.: Dynamical kernel of the Aladin-NH spectral limited-area model: Revised formulation and sensitivity experiments, Q. J. Roy. Meteor. Soc., 136, 155–169, https://doi.org/10.1002/qj.522, 2010.

Betts, A. K. and Jakob, C.: Study of diurnal cycle of convective precipitation over Amazonia using a single column model, J. Geophys. Res.-Atmos., 107, ACL 25-1–ACL 25-13, https://doi.org/10.1029/2002JD002264, 2002.

Bonaventura, L.: An introduction to semi-Lagrangian methods for geophysical scale flows, Lecture Notes, ERCOFTAC Leonhard Euler Lectures, SAM-ETH Zurich, Zurich, 2004.

Bryan, G. H., Wyngaard, J. C., and Fritsch, J. M.: Resolution Requirements for the Simulation of Deep Moist Convection, Mon. Weather Rev., 131, 2394–2416, https://doi.org/10.1175/1520-0493(2003)131<2394:RRFTSO>2.0.CO;2, 2003.

Bubnová, R., Hello, G., Bénard, P., and Geleyn, J.-F.: Integration of the Fully Elastic Equations Cast in the Hydrostatic Pressure Terrain-Following Coordinate in the Framework of the ARPEGE/Aladin NWP System, Mon. Weather Rev., 123, 515–535, https://doi.org/10.1175/1520-0493(1995)123<0515:IOTFEE>2.0.CO;2, 1995.

Callies, J., Bühler, O., and Ferrari, R.: The dynamics of mesoscale winds in the upper troposphere and lower stratosphere, J. Atmos. Sci., 73, 4853–4872, https://doi.org/10.1175/JAS-D-16-0108.1, 2016.

Cho, J. Y., Zhu, Y., Newell, R. E., Anderson, B. E., Barrick, J. D., Gregory, G. L., Sachse, G. W., Carroll, M. A., and Albercook, G. M.: Horizontal wavenumber spectra of winds, temperature, and trace gases during the Pacific Exploratory Missions: 1. Climatology, J. Geophys. Res.-Atmos., 104, 5697–5716, https://doi.org/10.1029/98JD01825, 1999a.

Cho, J. Y. N., Newell, R. E., and Barrick, J. D.: Horizontal wavenumber spectra of winds, temperature, and trace gases during the Pacific Exploratory Missions: 2. Gravity waves, quasi-two-dimensional turbulence, and vortical modes, J. Geophys. Res., 104, 16297–16308, https://doi.org/10.1029/1999JD900068, 1999b.

Courant, R., Friedrichs, K., and Lewy, H.: Über die partiellen Differenzengleichungen der mathematischen Physik, Mathematis-

che Annalen, 100, 32–74, https://doi.org/10.1007/BF01448839, 1928.

Dai, A. and Trenberth, K. E.: The diurnal cycle and its depiction in the community climate system model, J. Climate, 17, 930–951, https://doi.org/10.1175/1520-0442(2004)017<0930:TDCAID>2.0.CO;2, 2004.

Daley, R.: The normal modes of the spherical non-hydrostatic equations with applications to the filtering of acoustic modes normal modes of the spherical non-hydrostatic equations with applications to the filterin, Tellus A, 40, 96–106, https://doi.org/10.3402/tellusa.v40i2.11785, 1988.

Dirmeyer, P. A., Cash, B. A., Kinter, J. L., Jung, T., Marx, L., Satoh, M., Stan, C., Tomita, H., Towers, P., Wedi, N., Achuthavarier, D., Adams, J. M., Altshuler, E. L., Huang, B., Jin, E. K., and Manganello, J.: Simulating the diurnal cycle of rainfall in global climate models: Resolution versus parameterization, Clim. Dynam., 39, 399–418, https://doi.org/10.1007/s00382-011-1127-9, 2012.

Doms, G. and Baldauf, M.: A Description of the Nonhydrostatic Regional COSMO-Model Part I : Dynamics and Numerics, https://doi.org/10.5676/DWD_pub/nwv/cosmo-doc_5.05_I, 2018.

Done, J., Davis, C. A., and Weisman, M.: The next generation of NWP: Explicit forecasts of convection using the weather research and forecasting (WRF) model, Atmos. Sci. Lett., 5, 110–117, https://doi.org/10.1002/asl.72, 2004.

Dudhia, J.: A nonhydrostatic version of the Penn State-NCAR mesoscale model: validation tests and simulation of an Atlantic cyclone and cold front, Mon. Weather Rev., 121, 1493–1513, https://doi.org/10.1175/1520-0493(1993)121<1493:ANVOTP>2.0.CO;2, 1993.

Dueben, P. D., Wedi, N., Saarinen, S., and Zeman, C.: Global simulations of the atmosphere at 1.45 km grid-spacing with the integrated forecasting system, J. Meteorol. Soc. Japan, 98, 551–572, https://doi.org/10.2151/jmsj.2020-016, 2020.

ECMWF: Part IV : Physical processes, in: IFS Documentation CY45R1, ECMWF, 4, https://doi.org/10.21957/4whwo8jw0, 2018.

Fosser, G., Khodayar, S., and Berg, P.: Benefit of convection permitting climate model simulations in the representation of convective precipitation, Clim. Dynam., 44, 45–60, https://doi.org/10.1007/s00382-014-2242-1, 2014.

Fuhrer, O., Osuna, C., Lapillonne, X., Gysi, T., Bianco, M., Arteaga, A., and Schulthess, T. C.: Towards a performance portable, architecture agnostic implementation strategy for weather and climate models, Supercomputing Frontiers and Innovations, 1, 44–61, https://doi.org/10.14529/jsfi140103, 2014.

Fuhrer, O., Chadha, T., Hoefler, T., Kwasniewski, G., Lapillonne, X., Leutwyler, D., Lüthi, D., Osuna, C., Schär, C., Schulthess, T. C., and Vogt, H.: Near-global climate simulation at 1 km resolution: establishing a performance baseline on 4888 GPUs with COSMO 5.0, Geosci. Model Dev., 11, 1665–1681, https://doi.org/10.5194/gmd-11-1665-2018, 2018.

Gao, X. and Meriwether, J. W.: Mesoscale spectral analysis of in situ horizontal and vertical wind measurements at 6 km, J. Geophys. Res.-Atmos., 103, 6397–6404, https://doi.org/10.1029/97JD03074, 1998.

Gelb, A. and Gleeson, J. P.: Spectral Viscosity for Shallow Water Equations in Spherical Geometry, Mon. Weather Rev., 129, 2346–2360, https://doi.org/10.1175/1520-0493(2001)129<2346:SVFSWE>2.0.CO;2, 2001.

Guichard, F., Petch, J. C., Redelsperger, J. L., Bechtold, P., Chaboureau, J. P., Cheinet, S., Grabowski, W., Grenier, H., Jones, C. G., Köhler, M., Piriou, J. M., Tailleux, R., and Tomasini, M.: Modelling the diurnal cycle of deep precipitating convection over land with cloud-resolving models and single-column models, Q. J. Roy. Meteor. Soc., 130 C, 3139–3172, https://doi.org/10.1256/qj.03.145, 2004.

Held, I. M. and Soden, B. J.: Robust responses of the Hydrological Cycle to Global Warming, J. Climate, 19, 5686–5699, https://doi.org/10.1175/JCLI3990.1, 2006.

Hogan, R., Ahlgrimm, M., Balsamo, G., Beljaars, A., Berrisford, P., Bozzo, A., Giuseppe, F. D., Forbes, R. M., Haiden, T., Lang, S., Mayer, M., Polichtchouk, I., Sandu, I., Vitart, F., and Wedi, N.: Radiation in numerical weather prediction, ECMWF Technical Memoranda, 816, 1–49, https://doi.org/10.21957/2bd5dkj8x, 2017.

Hohenegger, C., Brockhaus, P., and Schär, C.: Towards climate simulations at cloud-resolving scales, Meteorol. Z., 17, 383–394, https://doi.org/10.1127/0941-2948/2008/0303, 2008.

Houze, R. A. and Betts, A. K.: Convection in GATE, Rev. Geophys. Space Phys., 19, 541–576, https://doi.org/10.1029/RG019i004p00541, 1981.

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E. J., Sorooshian, S., Tan, J., and Xie, P.: NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG), Algorithm Theoretical Basis Document (ATBD) Version 06, available at: https://gpm.nasa.gov/sites/default/files/2019-05/IMERG_ATBD_V06.pdf (last access: 13 July 2021), 2019a.

Huffman, G. J., Stocker, E. F., Bolvin, D. T., Nelkin, E. J., and Jackson, T.: GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree × 0.1 degree V06, https://doi.org/10.5067/GPM/IMERG/3B-HH/06, 2019b.

IDAWEB: IDAWEB, available at: https://gate.meteoswiss.ch/idaweb, last access: 15 January 2021.

Jablonowski, C. and Williamson, D. L.: The Pros and Cons of Diffusion, Filters and Fixers in Atmospheric General Circulation Models BT – Numerical Techniques for Global Atmospheric Models, in: Numerical Techniques for Global Atmospheric Models, edited by: Lauritzen, P., Jablonowski, C., Taylor, M., and Nair, R., Springer Berlin Heidelberg, Berlin, Heidelberg, 381–493, https://doi.org/10.1007/978-3-642-11640-7_13, 2011.

Jeevanjee, N.: Vertical Velocity in the Gray Zone, J. Adv. Model. Earth Syst., 9, 2304–2316, https://doi.org/10.1002/2017MS001059, 2017.

Jung, T., Miller, M. J., Palmer, T. N., Towers, P., Wedi, N., Achuthavarier, D., Adams, J. M., Altshuler, E. L., Cash, B. A., Kinter III, J. L., Marx, L., Stan, C., and Hodges, K. I.: High-Resolution Global Climate Simulations with the ECMWF Model in Project Athena: Experimental Design, Model Climate, and Seasonal Forecast Skill, J. Climate, 25, 3155–3172, https://doi.org/10.1175/JCLI-D-11-00265.1, 2012.

Kato, T.: Hydrostatic and Non-hydrostatic Simulations of the 6 August 1993 Kagoshima Torrential Rain, J. Meteorol. Soc. Japan Ser. II, 74, 355–363, https://doi.org/10.2151/jmsj1965.74.3_355, 1996.

Kato, T.: Hydrostatic and non-hydrostatic simulations of moist convection: Review and further study, Meteorol. Atmos. Phys., 63, 39–51, https://doi.org/10.1007/BF01025363, 1997.

Kato, T. and Saito, K.: Hydrostatic and Non-Hydrostatic Simulations of Moist Convection: Applicability of the Hydrostatic Approximation to a High-Resolution Model, J. Meteorol. Soc. Japan, 73, 59–77, https://doi.org/10.2151/jmsj1965.73.1_59, 1995.

Kendon, E. J., Roberts, N. M., Senior, C. A., and Roberts, M. J.: Realism of rainfall in a very high-resolution regional climate model, J. Climate, 25, 5791–5806, https://doi.org/10.1175/JCLI-D-11-00562.1, 2012.

Kendon, E. J., Stratton, R. A., Tucker, S., Marsham, J. H., Berthou, S., Rowell, D. P., and Senior, C. A.: Enhanced future changes in wet and dry extremes over Africa at convection-permitting scale, Nat. Commun., 10, 1794, https://doi.org/10.1038/s41467-019-09776-9, 2019.

Kühnlein, C., Deconinck, W., Klein, R., Malardel, S., Piotrowski, Z. P., Smolarkiewicz, P. K., Szmelter, J., and Wedi, N. P.: FVM 1.0: a nonhydrostatic finite-volume dynamical core for the IFS, Geosci. Model Dev., 12, 651–676, https://doi.org/10.5194/gmd-12-651-2019, 2019.

Lafore, J. P., Stein, J., Asencio, N., Bougeault, P., Ducrocq, V., Duron, J., Fischer, C., Héreil, P., Mascart, P., Masson, V., Pinty, J. P., Redelsperger, J. L., Richard, E., and Vilà-Guerau de Arellano, J.: The Meso-NH Atmospheric Simulation System. Part I: adiabatic formulation and control simulations, Ann. Geophys., 16, 90–109, https://doi.org/10.1007/s00585-997-0090-6, 1998.

Langhans, W., Schmidli, J., and Schär, C.: Bulk Convergence of Cloud-Resolving Simulations of Moist Convection over Complex Terrain, J. Atmos. Sci., 69, 2207–2228, https://doi.org/10.1175/JAS-D-11-0252.1, 2012.

Lauritzen, P. H., Jablonowski, C., Taylor, M. A., and Nair, R. D.: Numerical techniques for global atmospheric models, vol. 80, Springer Science & Business Media, 2011.

Lebo, Z. J. and Morrison, H.: Effects of horizontal and vertical grid spacing on mixing in simulated squall lines and implications for convective strength and structure, Mon. Weather Rev., 143, 4355–4375, https://doi.org/10.1175/MWR-D-15-0154.1, 2015.

Legates, D. R. and Willmott, C. J.: Mean seasonal and spatial variability in gauge-corrected, global precipitation, Int. J. Climatol., 10, 111–127, https://doi.org/10.1002/joc.3370100202, 1990.

Leutwyler, D., Lüthi, D., Ban, N., Fuhrer, O., and Schär, C.: Evaluation of the convection-resolving climate modeling approach on continental scales, J. Geophys. Res., 122, 5237–5258, https://doi.org/10.1002/2016JD026013, 2017.

Lindborg, E.: Can the atmospheric kinetic energy spectrum be explained by two-dimensional turbulence?, J. Fluid Mech., 388, 259–288, https://doi.org/10.1017/S0022112099004851, 1999.

Liu, H.-L.: Quantifying gravity wave forcing using scale invariance, Nat. Commun., 10, 2605, https://doi.org/10.1038/s41467-019-10527-z, 2019.

Lott, F. and Miller, M. J.: A new subgrid-scale orographic drag parametrization: Its formulation and testing, Q. J. Roy. Meteor. Soc., 123, 101–127, https://doi.org/10.1256/smsqj.53703, 1997.

Malardel, S. and Ricard, D.: An alternative cell-averaged departure point reconstruction for pointwise semi-Lagrangian transport schemes, Q. J. Roy. Meteor. Soc., 141, 2114–2126, https://doi.org/10.1002/qj.2509, 2015.

Malardel, S. and Wedi, N. P.: How does subgrid-scale parametrization influence nonlinear spectral energy fluxes in global NWP models?, J. Geophys. Res.-Atmos., 121, 5395–5410, https://doi.org/10.1002/2015JD023970, 2016.

Malardel, S., Wedi, N., Deconinck, W., Diamantakis, M., Kühnlein, C., Mozdzynski, G., Hamrud, M., and Smolarkiewicz, P.: A new grid for the IFS, ECMWF Newsletter, 146, 23–28, https://doi.org/10.21957/zwdu9u5i, 2016.

Manabe, S., Smagorinsky, J., and Strickler, R. F.: Simulated Climatology of a General Circulation Model With a Hydrologic Cycle 1, Mon. Weather Rev., 93, 769–798, https://doi.org/10.1175/1520-0493(1965)093<0769:scoagc>2.3.co;2, 1965.

Mishra, S. K. and Sahany, S.: Effects of time step size on the simulation of tropical climate in NCAR-CAM3, Clim. Dynam., 37, 689–704, https://doi.org/10.1007/s00382-011-0994-4, 2011.

Miyamoto, Y., Kajikawa, Y., Yoshida, R., Yamaura, T., Yashiro, H., and Tomita, H.: Deep moist atmospheric convection in a subkilometer global simulation, Geophys. Res. Lett., 40, 4922–4926, https://doi.org/10.1002/grl.50944, 2013.

Nastrom, G. D. and Gage, K. S.: A Climatology of Atmospheric Wavenumber Spectra of Wind and Temperature Observed by Commercial Aircraft, J. Atmos. Sci., 42, 950–960, https://doi.org/10.1175/1520-0469(1985)042<0950:ACOAWS>2.0.CO;2, 1985.

Neumann, P., Düben, P., Adamidis, P., Bauer, P., Brück, M., Kornblueh, L., Klocke, D., Stevens, B., Wedi, N., and Biercamp, J.: Assessing the scales in numerical weather and climate predictions: Will exascale be the rescue?, Philos. T. Roy. Soc. A, 377, 20180148, https://doi.org/10.1098/rsta.2018.0148, 2019.

Orlanski, I.: The quasi-hydrostatic approximation., J. Atmos. Sci., 38, 572–582, https://doi.org/10.1175/1520-0469(1981)038<0572:TQHA>2.0.CO;2, 1981.

Panosetti, D., Schlemmer, L., and Schär, C.: Bulk convergence behavior of convection-resolving simulations of summertime deep convection over land, Clim. Dynam., 55, 215–234, https://doi.org/10.1007/s00382-018-4229-9, 2018.

Panosetti, D., Schlemmer, L., and Schär, C.: Bulk and structural convergence at convection-resolving scales in real-case simulations of summertime moist convection over land, Q. J. Roy. Meteor. Soc., 145, 1427–1443, 2019.

Paulat, M., Frei, C., Hagen, M., and Wernli, H.: A gridded dataset of hourly precipitation in Germany: Its construction, climatology and application, Meteorol. Z., 17, 719–732, https://doi.org/10.1127/0941-2948/2008/0332, 2008.

Pearson, K. J., Lister, G. M., Birch, C. E., Allan, R. P., Hogan, R. J., and Woolnough, S. J.: Modelling the diurnal cycle of tropical convection across the "grey zone", Q. J. Roy. Meteor. Soc., 140, 491–499, https://doi.org/10.1002/qj.2145, 2014.

Prein, A. F., Gobiet, A., Suklitsch, M., Truhetz, H., Awan, N. K., Keuler, K., and Georgievski, G.: Added value of convection permitting seasonal simulations, Clim. Dynam., 41, 2655–2677, https://doi.org/10.1007/s00382-013-1744-6, 2013.

Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., Keller, M., Tölle, M., Gutjahr, O., Feser, F., Brisson, E., Kollet, S., Schmidli, J., Van Lipzig, N. P., and Leung, R.: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges, Rev. Geophys., 53, 323–361, https://doi.org/10.1002/2014RG000475, 2015.

Pudykiewicz, J., Benoit, R., and Staniforth, A.: Preliminary results from a partial lrtap model based on an existing meteorological forecast model, Atmos.-Ocean, 23, 267–303, https://doi.org/10.1080/07055900.1985.9649229, 1985.

Ramsauer, T., Weiß, T., and Marzahn, P.: Comparison of the GPM IMERG final precipitation product to RADOLAN weather radar data over the topographically and climatically diverse Germany, Remote Sensing, 10, https://doi.org/10.3390/rs10122029, 2018.

Raschendorfer, M.: The new turbulence parameterization of LM, COSMO Newsletter, 1, 89–97, 2001.

Reinhardt, T. and Seifert, A.: A three-category ice scheme for LMK, COSMO Newsletter, 6, 115–120, 2006.

Ricard, D., Lac, C., Riette, S., Legrand, R., and Mary, A.: Kinetic energy spectra characteristics of two convection-permitting limited-area models AROME and meso-NH, Q. J. Roy. Meteor. Soc., 139, 1327–1341, https://doi.org/10.1002/qj.2025, 2013.

Richter, D.: Ergebnisse methodischer Untersuchungen zur Korrektur des systematischen Meßfehlers des Hellmann-Niederschlagsmessers, 194, Selbstverlag des Deutschen Wetterdienstes, Offenbach am Main, 1995.

Ritter, B. and Geleyn, J.-F.: A Comprehensive Radiation Scheme for Numerical Weather Prediction Models with Potential Applications in Climate Simulations, Mon. Weather Rev., 120, 303–325, https://doi.org/10.1175/1520-0493(1992)120<0303:ACRSFN>2.0.CO;2, 1992.

Rockel, B., Will, A., and Hense, A.: The regional climate model COSMO-CLM (CCLM), Meteorol. Z., 17, 347–348, https://doi.org/10.1127/0941-2948/2008/0309, 2008.

Romero, R., Doswell, C. A., and Riosalido, R.: Observations and fine-grid simulations of a convective outbreak in Northeastern Spain: Importance of diurnal forcing and convective cold pools, Mon. Weather Rev., 129, 2157–2182, https://doi.org/10.1175/1520-0493(2001)129<2157:OAFGSO>2.0.CO;2, 2001.

Ross, B. B. and Orlanski, I.: The Circulation Associated with a Cold Front. Part II: Moist Case, J. Atmos. Sci., 35, 445–465, https://doi.org/10.1175/1520-0469(1978)035<0445:tcawac>2.0.co;2, 1978.

Schär, C., Fuhrer, O., Arteaga, A., Ban, N., Charpilloz, C., Girolamo, S. D., Hentgen, L., Hoefler, T., Lapillonne, X., Leutwyler, D., Osterried, K., Panosetti, D., Rüdisühli, S., Schlemmer, L., Schulthess, T. C., Sprenger, M., Ubbiali, S., and Wernli, H.: Kilometer-Scale Climate Models, B. Am. Meteorol. Soc., 101, E567–E587, https://doi.org/10.1175/BAMS-D-18-0167.1, 2020.

Schlemmer, L., Schär, C., Lüthi, D., and Strebel, L.: A Groundwater and Runoff Formulation for Weather and Climate Models, J. Adv. Model. Earth Syst., 10, 1809–1832, https://doi.org/10.1029/2017MS001260, 2018.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., and Ziese, M.: GPCC Monitoring Product: Near Real-Time Monthly Land-Surface Precipitation from Rain-Gauges based on SYNOP and CLIMAT data, Deutscher Wetterdienst, https://doi.org/10.5676/DWD_GPCC/MP_M_V6_100, 2018.

Schulthess, T. C., Bauer, P., Wedi, N., Fuhrer, O., Hoefler, T., and Schär, C.: Reflecting on the Goal and Baseline for Exascale Computing: A Roadmap Based on Weather and Climate Simulations, Comput. Sci. Eng., 21, 30–41, https://doi.org/10.1109/MCSE.2018.2888788, 2019.

Schumann, U.: The horizontal spectrum of vertical velocities near the tropopause from global to gravity wave scales, J. Atmos. Sci., 76, 3847–3862, https://doi.org/10.1175/JAS-D-19-0160.1, 2019.

Sevruk, B.: Systematischer Niederschlagsmessfehler in der Schweiz, in: Der Niederschlag in der Schweiz, no. 31 in Beiträge zur Geologie der Schweiz-Hydrologie, chap. 3.1, pp. 65–75, Schweizerische Geotechnische Kommission, Zürich, 1985.

Sevruk, B.: Rainfall Measurement: Gauges, in: Encyclopedia of Hydrological Sciences, Part 4. Hydrometeorology, edited by: Anderson, M. G., John Wiley & Sons, Ltd., 35, 8, https://doi.org/10.1002/0470848944.hsa038, 2005.

Skamarock, W. C.: Evaluating Mesoscale NWP Models Using Kinetic Energy Spectra, Mon. Weather Rev., 132, 3019–3032, https://doi.org/10.1175/MWR2830.1, 2004.

Skamarock, W. C. and Klemp, J. B.: A time-split nonhydrostatic atmospheric model for weather research and forecasting applications, J. Comput. Phys., 227, 3465–3485, https://doi.org/10.1016/j.jcp.2007.01.037, 2008.

Smagorinsky, J.: General circulation experiments with the primitive equations, Mon. Weather Rev., 91, 99–164, https://doi.org/10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2, 1963.

Smolarkiewicz, P. K. and Pudykiewicz, J. A.: A Class of Semi-Lagrangian Approximations for Fluids, J. Atmos. Sci., 49, 2082–2096, https://doi.org/10.1175/1520-0469(1992)049<2082:ACOSLA>2.0.CO;2, 1992.

Staniforth, A. and Côté, J.: Semi-Lagrangian Integration Schemes for Atmospheric Models – A Review, Mon. Weather Rev., 119, 2206–2223, https://doi.org/10.1175/1520-0493(1991)119<2206:SLISFA>2.0.CO;2, 1991.

Stephens, G. L., L'Ecuyer, T., Forbes, R., Gettlemen, A., Golaz, J. C., Bodas-Salcedo, A., Suzuki, K., Gabriel, P., and Haynes, J.: Dreary state of precipitation in global models, J. Geophys. Res.-Atmos., 115, 1–14, https://doi.org/10.1029/2010JD014532, 2010.

Sun, Y., Solomon, S., Dai, A., and Portmann, R. W.: How often does it rain?, J. Climate, 19, 916–934, https://doi.org/10.1175/JCLI3672.1, 2006.

Tiedtke, M.: A comprehensive mass flux scheme for cumulus parameterization in large-scale models, Mon. Weather Rev., 117, 1779–1800, https://doi.org/10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2, 1989.

Ubbiali, S., Schär, C., Schlemmer, L., and Schulthess, T. C.: A numerical analysis of six physics-dynamics coupling schemes for atmospheric models, J. Adv. Model. Earth Sy., in review, 2021.

Vergara-Temprado, J., Ban, N., Panosetti, D., Schlemmer, L., and Schär, C.: Climate models permit convection at much coarser resolutions than previously considered, J. Climate, 33, 1915–1933, https://doi.org/10.1175/JCLI-D-19-0286.1., 2020.

Vergara-Temprado, J., Ban, N., and Schär, C.: Extreme Sub-Hourly Precipitation Intensities Scale Close to the Clausius-Clapeyron Rate Over Europe, Geophys. Res. Lett., 48, e2020GL089506, https://doi.org/10.1029/2020GL089506, 2021.

Villarini, G. and Krajewski, W. F.: Review of the different sources of uncertainty in single polarization radar-based estimates of rainfall, Surv. Geophys., 31, 107–129, https://doi.org/10.1007/s10712-009-9079-x, 2010.

Wang, S., Liu, J., Wang, J., Qiao, X., and Zhang, J.: Evaluation of GPM IMERG V05B and TRMM 3B42V7 Precipitation products over high mountainous tributaries in Lhasa with dense rain gauges, Remote Sensing, 11, https://doi.org/10.3390/rs11182080, 2019.

Wedi, N., Yessad, K., and Untch, A.: The non-hydrostatic global IFS/ARPEGE model: model formulation and testing, ECMWF Technical Memoranda, p. 34, https://doi.org/10.21957/tl4f0ao4t, 2009.

Wedi, N. P.: Increasing horizontal resolution in numerical weather prediction and climate simulations: Illusion or panacea?, Philos. T. Roy. Soc. A, 372, 20130289, https://doi.org/10.1098/rsta.2013.0289, 2014.

Wedi, N. P., Hamrud, M., and Mozdzynski, G.: A fast spherical harmonics transform for global NWP and climate models, Mon. Weather Rev., 141, 3450–3461, https://doi.org/10.1175/MWR-D-13-00016.1, 2013.

Wedi, N. P., Dueben, P., Anantharaj, V. G., Bauer, P., Boussetta, S., Browne, P., Deconinck, W., Gaudin, W., Hadade, I., Hatfield, S., Iffrig, O., Lopez, P., Maciel, P., Mueller, A., Polichtchouk, I., Saarinen, S., Quintino, T., and Vitart, F.: A baseline for global weather and climate simulations at 1 km resolution, J. Adv. Model. Earth Sy., 12, e2020MS002192, https://doi.org/10.1029/2020MS002192, 2020.

Weisman, M. L., Skamarock, W. C., and Klemp, J. B.: The Resolution Dependence of Explicitly Modeled Convective Systems, Mon. Weather Rev., 125, 527–548, https://doi.org/10.1175/1520-0493(1997)125<0527:TRDOEM>2.0.CO;2, 1997.

Wicker, L. J. and Skamarock, W. C.: Time-Splitting Methods for Elastic Models Using Forward Time Schemes, Mon. Weather Rev., 130, 2088–2097, https://doi.org/10.1175/1520-0493(2002)130<2088:TSMFEM>2.0.CO;2, 2002.

Williamson, D. L. and Olson, J. G.: Dependence of aqua-planet simulations on time step, Q. J. Roy. Meteor. Soc., 129, 2049–2064, https://doi.org/10.1256/qj.02.62, 2003.

Winterrath, T., Brendel, C., Hafer, M., Junghänel, T., Klameth, A., Walawender, E., Weigl, E., and Becker, A.: Erstellung einer radargestützten Niederschlagsklimatologie, 251, Deutscher Wetterdienst, https://doi.org/10.17169/refubium-25153, 2017.

Wüest, M., Frei, C., Altenhoff, A., Hagen, M., Litschi, M., and Schär, C.: A gridded hourly precipitation dataset for Switzerland using rain-gauge analysis and radar-based disaggregation, Int. J. Climatol., 30, 1764–1775, https://doi.org/10.1002/joc.2025, 2010.

Yang, G. Y. and Slingo, J.: The diurnal cycle in the tropics, Mon. Weather Rev., 129, 784–801, https://doi.org/10.1175/1520-0493(2001)129<0784:TDCITT>2.0.CO;2, 2001.

Zeman, C., Wedi, N. P., Dueben, P. D., Ban, N., and Schär, C.: Model intercomparison of COSMO 5.0 and IFS 45r1 at kilometer-scale grid spacing (Version 1.0) [data set], Zenodo, https://doi.org/10.5281/zenodo.4479130, 2021.