Table 1. Model description and setup.

	COSMO	IFS
Numerics:	Split Explicit	Semi-Lagrangian and Semi-Implicit
Vertical Velocity:	Nonhydrostatic	Hydrostatic
Horizontal Discretization:	Rotated Lat/Lon	Spectral and Reduced Gaussian, octahedral
Resolution Setups:	$\Delta x = 12\,\text{km}$ (0.11°), $\Delta t \in \{90\,\text{s}, 40\,\text{s}\}$	$\Delta x = 9\,\text{km}$ (TCo 1279), $\Delta t \in \{450\,\text{s}, 240\,\text{s}\}$
	$\Delta x = 4.4\,\text{km}$ (0.04°), $\Delta t \in \{40\,\text{s}, 20\,\text{s}\}$	$\Delta x = 4.5\,\text{km}$ (TCo 2559), $\Delta t \in \{240\,\text{s}, 120\,\text{s}\}$
	$\Delta x = 2.2\,\text{km}$ (0.02°), $\Delta t \in \{20\,\text{s}, 10\,\text{s}\}$	$\Delta x = 2.9\,\text{km}$ (TCo 3999), $\Delta t \in \{120\,\text{s}, 60\,\text{s}\}$
Vertical Levels:	60 (up to $\sim 23.5\,\text{km}$)[c]	137 (up to $\sim 80.5\,\text{km}$)[c]
Convection Param.:	Deep Convection off[a]	Deep Convection off[b]
	Shallow Convection on	Shallow Convection on

[a]For $\Delta x = 12\,\text{km}$, $\Delta t = 90\,\text{s}$ also one run with deep convection parameterization on

[b]For $\Delta x = 9\,\text{km}$, $\Delta t = 450\,\text{s}$ also one run with deep convection parameterization on

[c]The vertical spacing is the same for all resolution setups.

whereas COSMO is run regionally with the ECMWF operational analysis data as lateral boundary conditions on a limited area domain ranging from 361×361 ($\Delta x = 12\,\text{km}$) to 1542×1542 ($\Delta x = 2.2\,\text{km}$) 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 $\Delta x = 12\,\text{km}$, 4.4 km, and 2.2 km, whereas the grid spacings used with IFS approximately correspond to $\Delta x = 9\,\text{km}$, 4.5 km, and 2.9 km. Both, COSMO and IFS usually use deep convection parameterization for coarser resolutions such as $\Delta x = 12\,\text{km}$ for COSMO or $\Delta x = 9\,\text{km}$ 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 timestep in the respective simulations, each horizontal resolution setup with explicit deep convection has also been run with a smaller than usual timestep. 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 ~~horizontal diffusion~~ diffusion from a monotonic 4th-order linear scheme with orographic limiter. This will give us some idea about the influence of ~~horizontal~~ diffusion on the model results and might explain some characteristic differences between IFS and COSMO. In COSMO, 4th-order diffusion is applied on model levels by introducing an additional operator at the right hand side of the prognostic equation, similar to

$$\frac{\partial \psi}{\partial t} = S(\psi) + (-1)^{4/3}\alpha_4 \mathtt{diff} \cdot c_d \cdot \nabla^4 \psi\,, \tag{1}$$

where ψ is the prognostic variable~~and~~ , S represents all physical and dynamical source terms for ψ~~. The prognostic variables on which horizontal diffusion is applied are wind, temperature, pressure, specific humidity, and cloud water content. The~~ , c_d is the default diffusion coefficient ~~is α_4 is dependent on the horizontal and temporal resolution such that $\alpha_4 = (\Delta x/2)^4/\Delta t$. This coefficient can be multiplied with a factor, which we will hereafter call~~ in the model, and `diff` ~~,~~ is the factor that can be set in order to ~~apply more or less smoothing to the mentioned variables~~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 `diff` = 1 means that the diffusion coefficient ~~remains unchanged and corresponds to the default value α_4~~c_d remains unchanged. Any value of `diff` smaller than one decreases the explicit diffusion coefficient and any value larger than one increases explicit 4th-order diffusion strength. In our default setup for the intercomparison, COSMO has been run with no explicit 4th-order linear horizontal diffusion, which means `diff` was set to zero. For this experiment, the 2.2 km setup with a timestep of $\Delta t = 20\,\mathrm{s}$ has been used, but with numbers for `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) and 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 $0.1^\circ \times 0.1^\circ$ 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 consists 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 stems from the Global Precipitation Climate Centre (GPCC) which is operated by the German Weather Service (DWD, Deutscher Wetter Dienst). 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 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.