Water isotopic characterisation of the cloud-circulation coupling in the North Atlantic trades. Part 1: A process-oriented evaluation of COSMO$_{iso}$ simulations with EUREC[4]A observations

Leonie Villiger[1,2], Marina Dütsch[3], Sandrine Bony[4], Marie Lothon[5], Stephan Pfahl[6], Heini Wernli[1], Pierre-Etienne Brilouet[7], Patrick Chazette[8], Pierre Coutris[9], Julien Delanoë[10], Cyrille Flamant[11], Alfons Schwarzenboeck[9], Martin Werner[12], and Franziska Aemisegger[1]

[1]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland
[2]Institute for Environmental Decisions, ETH Zurich, Zurich, Switzerland
[3]Department of Meteorology and Geophysics, University of Vienna, Vienna, Austria
[4]LMD/IPSL, Sorbonne Université, CNRS, Paris, France
[5]Laboratoire d'Aérologie, University of Toulouse, CNRS, Toulouse, France
[6]Institute of Meteorology, Freie Universität Berlin, Berlin, Germany
[7]CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France
[8]LSCE/IPSL, CNRS-CEA-UVSQ, University Paris-Saclay, Gif sur Yvette, France
[9]LAMP, Université Clermont Auvergne, CNRS, Clermont-Ferrand, France
[10]LATMOS/IPSL, Université Paris-Saclay, UVSQ, Guyancourt, France
[11]LATMOS/IPSL, Sorbonne Université, CNRS, Paris, France
[12]Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany

Correspondence: Leonie Villiger (leonie.villiger@env.ethz.ch)

Abstract. Naturally available, stable and heavy water molecules such as HDO and $H_2^{18}O$ have a lower saturation vapour pressure than the most abundant light water molecule $H_2^{16}O$ and therefore these heavy water molecules preferentially condense and rain out during cloud formation. Stable water isotope observations thus have the potential to provide information on cloud processes in the trade-wind region, in particular when combined with high-resolution model simulations. In order to evaluate this potential, nested COSMO$_{iso}$ simulations with explicit convection and horizontal grid spacings of 10, 5, and 1 km were carried out in this study over the tropical Atlantic for the time period of the EUREC[4]A field experiment. The comparison to airborne in situ and remote sensing observations shows that the three simulations are able to distinguish between different mesoscale cloud organisation patterns as well as between periods with comparatively high and low rain rates. Cloud fraction and liquid water content show a better agreement with aircraft observations with higher spatial resolution because they show strong spatial variations on the scale of a few kilometres. A low-level cold-dry bias, including too depleted vapour in the subcloud and cloud layer and too enriched vapour in the free troposphere, is found in all three simulations. Furthermore, the simulated secondary isotope variable d-excess in vapour is overestimated compared to observations. Special attention is given to the cloud base level, the formation altitude of shallow cumulus clouds. The temporal variability of the simulated isotope variables at cloud base agrees reasonably well with observations, with correlations of the flight-to-flight data as high as 0.7 for δ^2H and d-excess. A close examination of isotopic characteristics under precipitating clouds, non-precipitating clouds, clear-sky and dry-warm patches at the altitude of cloud base shows that these different environments are represented faithfully in

the model with similar frequency of occurrence, isotope signals and specific humidity anomalies as found in the observations. Furthermore, it is shown that the δ^2H of cloud base vapour at the hourly timescale is mainly controlled by mesoscale transport and not by local microphysical processes, while the d-excess is mainly controlled by large-scale drivers. Overall, this evaluation of COSMO$_{iso}$, including the isotopic characterisation of different cloud base environments, suggests that the simulations can be used for investigating the role of atmospheric circulations on different scales for controlling the formation of shallow cumulus clouds in the trade-wind region, as will be done in part 2 of this study.

1 Introduction

Shallow clouds, ubiquitous in the trade-wind region, substantially contribute to the cooling of the Earth's climate through their shortwave radiative effect. Their response to climate change is unclear, contributing to a large part of the uncertainty of climate projections (e.g., Bony and Dufresne, 2005; Zelinka et al., 2017; Schneider et al., 2017). The cloud fraction at cloud base, in particular, has been identified as a key variable for the spread of the modelled feedback of these clouds to climate change (Bony et al., 2017). Therefore, understanding the processes controlling the variability of cloudiness at cloud base is of utmost importance.

The field campaign EUREC[4]A (Elucidating the role of clouds-circulation coupling in climate; Stevens et al., 2021), which took place in early 2020 in the vicinity of the Caribbean island Barbados, was designed to provide observational constraints on the mechanisms that control shallow trade-wind clouds. A focus of the campaign was the mesoscale organisation pattern of these clouds, with the most frequent four patterns named Sugar, Gravel, Flower, and Fish (Stevens et al., 2020; Bony et al., 2020; Schulz, 2022). As part of EUREC[4]A, the EUREC[4]A$_{iso}$ component coordinated a multi-platform network of stable water isotope measurements (Bailey et al., 2023). Stable water isotopologues (hereafter simply named stable water isotopes) reflect the integral of moist processes experienced by an air parcel during transport (Gat, 1996; Galewsky et al., 2016). Hence, they are a promising tool for bridging the gap between microphysical processes at the scale of clouds and transport processes on larger scales.

Here, we use the isotope measurements that were conducted onboard the French aircraft SAFIRE ATR-42 (hereafter ATR; Bony et al., 2022), whose mission during EUREC[4]A was to provide a detailed characterisation of atmospheric properties near the cloud-base level and within the subcloud layer. The measurements, which are limited in space and time, are complemented with three numerical simulations that were performed using the isotope-enabled version of the non-hydrostatic weather forecast model from the Consortium for Small Scale Modelling COSMO$_{iso}$ (Steppeler et al., 2003; Pfahl et al., 2012). Providing a description of the three-dimensional distribution of isotopes at hourly intervals and at high spatial resolution, the simulations can be used to investigate the impact of the atmospheric circulation and the physical processes embedded in the flow on isotope signals. Specifically, we test the hypothesis that isotopes are modulated by both microphysical processes in the cloud-relative overturning circulation and variations in the large-scale flow. In turn, this means that isotopes have the potential to provide an observational constraint on these processes that are otherwise very difficult to observe. This hypothesis can only be addressed with multi-scale, regional model simulations with realistic, time-dependent boundary conditions, which provide a

statistically robust, three-dimensional picture of cloud characteristics and associated isotope signals. In addition, the EUREC[4]A ATR observational isotope dataset is essential for evaluating the COSMO$_{iso}$ simulation data before using it to deepen our process understanding. For this evaluation, three objectives are pursued: (1) evaluate the ability of COSMO$_{iso}$ to reproduce key characteristics of shallow cumulus clouds in the trade-wind region, (2) assess the effect of the spatial model resolution on different variables characterising the occurrence and the environment of shallow trade-wind cumulus clouds, and (3) conduct a feature-based assessment (including cloudy and clear-sky patches) of isotope variability at cloud base, where shallow cumulus cloud formation is initiated.

This section continues with a short introduction into stable water isotope physics, a general description of the trade-wind boundary layer, and a summary of the results from earlier, relevant model evaluation studies. After the introduction (Sect. 1), the applied datasets and methods are described (Sect. 2), followed by the results of the evaluation of the three COSMO$_{iso}$ simulations (Sect. 3) and the assessment of the isotope variability at cloud base (Sect. 4). The paper ends with a summary of the most important findings and their consequences for investigating the role of mesoscale circulations for shallow cumulus cloud formation using water isotopes as tracers (Sect. 5).

1.1 Stable water isotopes

Heavy stable water isotopes are water molecules containing a heavy hydrogen (i.e., ^{1}H^{2}H^{16}O also referred to as HDO) or oxygen atom (i.e., ^{1}H$_{2}$^{18}O). As a result, they have lower saturation vapour pressures and lower diffusion velocities than their light counterpart (i.e., ^{1}H$_{2}$^{16}O), which leads to a change in the relative abundance of heavy to light isotopes during phase transitions. The isotopic composition of a water sample is typically communicated with the so-called δ-notation for ^{2}H and ^{18}O, respectively:

$$\delta^{2}\text{H [‰]} = \left(\frac{^{2}R_{sample}}{^{2}R_{VSMOW}} - 1 \right) \cdot 1000 \tag{1}$$

$$\delta^{18}\text{O [‰]} = \left(\frac{^{18}R_{sample}}{^{18}R_{VSMOW}} - 1 \right) \cdot 1000 \tag{2}$$

The R in Eq. 1 and Eq. 2 stands for the atomic ratio of the concentration of the heavy (rare) to the light (most abundant) isotope, i.e., $^{2}R = \frac{[^{2}\text{H}]}{[^{1}\text{H}]}$ and $^{18}R = \frac{[^{18}\text{O}]}{[^{16}\text{O}]}$, in the water sample and the internationally accepted Vienna Mean Ocean Standard Water 2 (VSMOW; International Atomic Energy Agency, 2017). Samples with high δ^{2}H and δ^{18}O values are referred to as enriched (in heavy isotopes), and samples with low δ^{2}H and δ^{18}O values are referred to as depleted (in heavy isotopes). The relative variations of δ^{2}H and δ^{18}O are assessed with the deuterium excess (d-excess), defined as $d = \delta^{2}\text{H} - 8 \times \delta^{18}\text{O}$, which is a measure of the thermodynamic disequilibrium of the environment during phase transitions (e.g., Gat, 1996; Pfahl and Wernli, 2008).

The usefulness of isotopes has been demonstrated in numerous earlier studies. For example, it is known that extratropical cyclones and cold front passages leave a clear and attributable signal in surface isotope measurements (e.g., Aemisegger et al., 2015; Aemisegger, 2018; Graf et al., 2019; Thurnherr et al., 2020). In the (sub)tropics, isotope observations have been linked

to the large-scale circulation (Torri et al., 2017; Guilpart et al., 2017), the daily cycle of the boundary layer (Noone et al., 2011), and precipitation efficiency in shallow convection (Bailey et al., 2015). The links identified in these studies build on the facts that (1) the isotope signal of water vapour is conserved during transport and modified during phase changes and therefore reflects pathway-specific processes, and (2) that the formation altitude of rainwater largely determines its isotopic composition. Airborne measurements have been considered highly valuable as they provide a three-dimensional description of isotope gradients (Dyroff et al., 2015; Sodemann et al., 2017; Salmon et al., 2019; Chazette et al., 2021), which, for instance, allows investigating vertical mixing processes. Furthermore, isotope-enabled models have been shown to be useful in looking into tropical deep convection (Bony et al., 2008; Risi et al., 2008; Blossey et al., 2010; Risi et al., 2020, 2021; de Vries et al., 2022), entrainment of free-tropospheric air into the tropical marine boundary layer (Risi et al., 2019), or precipitation-driven cold pool dynamics (Torri and Kuang, 2016a, b; Torri, 2021).

1.2 The trade-wind boundary layer

In this study, we focus on the isotopic characteristics of the lower troposphere in the trade-wind region, which is typically divided into the subcloud layer, the cloud layer, and the free troposphere (Fig. 1; see also Wood, 2012, in which many of the relevant processes described here are addressed in more detail). The subcloud layer and the cloud layer make up the boundary layer. The subcloud layer connects the cloud layer with the ocean surface, providing the buoyancy and moisture for the formation of new clouds. It is often assumed to be a well-mixed layer with negligible vertical gradients in humidity and temperature. Recently, however, Albright et al. (2022, their Fig. 2 and 4) showed that the well-mixed layer (its top being characterised by a jump in specific humidity, potential temperature, or relative humidity) does not always extend to the subcloud layer top (identified via buoyancy measures and corresponding to the lifting condensation level in measurements), but is sometimes separated from it by a thin transition layer. Furthermore, they found that the subcloud layer depth and fluxes are largely controlled by near-surface wind speed, while moisture and heat within the subcloud layer are more sensitive to the thermodynamic conditions above the subcloud layer top. Note that the occurrence of cold pools can modify this picture, generating a smaller, cooler mixed layer overlaid by a stable layer (Touzé-Peiffer et al., 2022).

Directly adjunct to the subcloud layer top is the cloud base level, where clouds start to form and may grow vertically until reaching the trade inversion, which represents the upper bound of the cloud layer. The horizontal extent of the clouds is typically largest at the cloud base, but, if the clouds develop up to the trade inversion level, a secondary maximum may emerge at their top in the form of stratiform outflows. This leads to a bi-modal vertical distribution of cloud fraction (Vial et al., 2019, e.g.,). The contrasting radiative properties of the cloudy and clear-sky environments in the cloud layer can induce shallow circulations (Naumann et al., 2017), which in turn influence the properties of the subcloud layer below. Another process through which the cloud layer feeds back to the subcloud layer is the formation of evaporatively driven cold pools, which create cold-dry density currents that spread radially at the surface. They may initiate new convection at their edges through mechanical lifting and buoyancy caused by increased moisture contents (e.g., Tompkins, 2001; Schlemmer and Hohenegger, 2014; Torri et al., 2015; Vogel et al., 2021).

The trade inversion that caps the cloud layer forms due to the large-scale subsidence typically found in the free troposphere of the trade-wind region. This large-scale subsidence corresponds to the descending branch of the Hadley cell. The subsidence rate is determined by the balance between adiabatic warming and radiative cooling and is roughly $35\,\mathrm{hPa\,d^{-1}}$ (Salathé and Hartmann, 1997; Holton and Hakim, 2013). Occasionally, however, the Hadley-cell-like descent is replaced by air masses descending much faster ($\sim 200\,\mathrm{hPa\,d^{-1}}$) from the extratropics along the slanted isentropes (Aemisegger et al., 2021b; Villiger et al., 2022).

Each of the above-described layers (Fig. 1) is characterised by a distinct isotopic composition of vapour. Generally, the vapour is more enriched in heavy isotopes closer to the surface. With increasing altitude, as soon as condensation and rainout occur, the vapour continuously loses its heavy isotopes and becomes lighter. The d-excess is assumed to be high near the ocean surface and to decrease with altitude due to a slight temperature dependency (Thurnherr et al., 2021). Once very dry conditions are reached (at about $12\,\mathrm{km}$ and above; this altitude was identified in vertical cross-sections [not shown] of the data from the numerical simulations described in Sect. 2) the non-linearity of the δ-scale causes a rapid increase of the d-excess relative to the values observed near the surface (Dütsch et al., 2017). The large-scale circulation can substantially alter the isotopic composition (Aemisegger et al., 2021b) by influencing the conditions under which the vapour is evaporated from the ocean surface or by advecting air masses from other regions (e.g., more depleted air from higher latitudes).

At cloud base, different processes are at play shaping the isotope variability. On the one hand, transport within mesoscale systems accumulates freshly evaporated and enriched vapour in areas where clouds are being formed and brings dry and depleted air from high altitudes to cloud base into the clear-sky environments between clouds (radiatively driven subsidence). The d-excess is expected to be high in convective updrafts bearing the signal of ocean evaporation and comparably lower for moisture in clear-sky environments, which has previously gone through cloud formation processes (i.e., moisture detrained from clouds). Note that in the shallow trade-wind cumuli regime, the air does not become dry enough for the non-linearity of the δ-scale to affect the d-excess. On the other hand, microphysical processes, such as the formation of cloud droplets or the partial or full evaporation of cloud and rain droplets, impact the cloud vapour's isotope signal. If the formed condensate is removed from the cloud through precipitation, the cloud layer's total water gets depleted in heavy isotopes. The vapour in clear-sky environments is indirectly influenced by this process, as ultimately, the cloudy air gets detrained into clear-sky environments and subsides back to cloud base. Both processes, mesoscale transport and microphysics, taken together, thus, lead to contrasting isotope signals in cloudy and clear-sky environments at cloud base.

1.3 Earlier model evaluation studies

Several studies have evaluated the performance of models in the lower troposphere of the trade-wind region during EUREC[4]A. Beucher et al. (2022) tested the oversea configuration of the French regional model AROME, called AROME-OM (its output used by Dauhut et al., 2022, as initial and lateral boundary conditions for large eddy simulations). In their model setup, deep convection is treated explicitly, but shallow convection is parameterised. Beucher et al. (2022) found that the AROME-OM produces a too deep subcloud layer associated with a dry bias ($-1\,\mathrm{g\,kg^{-1}}$ in specific humidity) and has a cold bias in the cloud layer ($-0.5\,\mathrm{K}$ in potential temperature). Further, the AROME-OM is able to produce stratiform clouds near the trade inversion,

but their occurrence is underestimated. Besides identifying these shortcomings, Beucher et al. (2022) demonstrated the capacity
of the AROME-OM to predict the different mesoscale cloud organisation patterns and their associated environment. Savazzi
et al. (2022) looked into lower-tropospheric biases of meridional winds in Integrated Forecasting System (IFS) forecasts and
ERA5 reanalyses. They found a weak wind speed bias at altitudes below 5 km during local day and a strong wind speed bias
below 2 km during local nighttime. Here, a further model evaluation in the framework of EUREC^{4}A is added, with a focus on
isotopes.

Since we are using COSMO$_{iso}$ simulations, we summarise the most relevant findings from earlier model evaluation studies
of COSMO and COSMO$_{iso}$ used at the km-scale resolution, albeit conducted mostly in regions other than the trades. Ban
et al. (2014) showed that turning off the deep convection scheme leads to a more accurate diel cycle of surface temperature
and precipitation over Europe. Vergara-Temprado et al. (2020) performed COSMO simulations over Europe at seven differ-
ent horizontal grid spacings, ranging from 50 km to 2.2 km. At each resolution, Vergara-Temprado et al. (2020) carried out
one simulation with the full Tiedtke (1989) scheme, one with a parameterisation of shallow convection only, and another one
without any convection parameterisation. They found that at grid spacings of 25 km or smaller, explicit convection improves
the model skill with regard to surface temperature, precipitation, and top-of-atmosphere radiation. The improvement over the
simulations with fully parameterised convection applies to both simulations with partly resolved (deep convection) and fully
resolved convection (deep and shallow). Furthermore, the study showed that the representation of some variables strongly
depends on the spatial resolution of the model, while for others, the convection parametrisation settings (i.e., resolved vs. pa-
rameterised) are crucial. For example, hourly precipitation and outgoing longwave radiation are mostly determined by whether
the convection is resolved or parameterised, while net shortwave radiation is more sensitive to the spatial resolution (due to its
strong dependence on cloud fraction). The grid-spacing sensitivity of COSMO was further investigated by Heim et al. (2021).
They carried out five convection-resolving simulations over the tropical Atlantic with a grid spacing ranging from 12 km to
500 m and found, for instance, that the low-cloud fraction decreases with finer resolution.

The isotope-enabled version of COSMO has been used and evaluated previously in different regions, but never in the tropical
trade-wind region. Thurnherr et al. (2021) used COSMO$_{iso}$ simulations with explicit convection over the Southern Ocean and
compared them to ship-based and radiosonde measurements. They observed that the vapour in the lowest model level was
too depleted and had too high d-excess values compared to in situ measurements (their Fig. 9). They argued that too strong
vertical mixing possibly caused the observed bias and that the offset in d-excess could be reduced by using the smooth regime
of Merlivat and Jouzel (1979)'s formulation for non-equilibrium fractionation instead of the formulation by Pfahl and Wernli
(2009) developed with observations from the Mediterranean. Previous studies (e.g., Steen-Larsen et al., 2015; Aemisegger
and Sjolte, 2018; Bonne et al., 2019) showed that using the non-equilibrium fractionation factor corresponding to the smooth
regime in the Merlivat and Jouzel (1979) formulation provides the most reliable results.

Dahinden et al. (2021) and de Vries et al. (2022) evaluated COSMO$_{iso}$'s humidity and isotope signals in the free troposphere
in the context of the West African monsoon. Dahinden et al. (2021), who used a setup with resolved convection in the vicinity
of the Canary Islands, found that COSMO$_{iso}$ is overly moist and enriched at altitudes ≥ 6 km compared to in situ aircraft
observations. However, the day-to-day variability of COSMO$_{iso}$'s humidity and δ^2H agreed well with ground-based remote

sensing observations from Teneriffe (FTIR). de Vries et al. (2022) performed three COSMO$_{iso}$ simulations in the West African monsoon region and tested the effect of parameterised versus explicit convection and different spatial resolutions. Independent of the modelling setup, COSMO$_{iso}$ produced a distinct bias towards too high δ^2H at 4.2 km compared to satellite-based remote sensing observations from IASI. Despite the offset, the temporal evolution of the simulations agreed well with the observations, lending confidence to the model's ability to reproduce the mesoscale to synoptic-scale variability in the water vapour isotope fields.

The literature summarised above suggests that simulations should be done by using an explicit representation of convection and for variables such as the cloud fraction a fine spatial resolution should be used (Ban et al., 2014; Vergara-Temprado et al., 2020; Heim et al., 2021), albeit without expecting a significant improvement for water vapour isotope variables (de Vries et al., 2022). Furthermore, we anticipate a depletion bias in the near-surface vapour (Thurnherr et al., 2021) and an enriched bias in the free-tropospheric vapour but overall a good performance in terms of the mesoscale to synoptic-scale variability in humidity and isotope variables (Dahinden et al., 2021; de Vries et al., 2022). Motivated by Thurnherr et al. (2021)'s consideration, we used the smooth regime of Merlivat and Jouzel (1979) in our simulations to formulate non-equilibrium fractionation during evaporation from the ocean surface. Furthermore, we applied a nested model setup, because earlier isotope-enabled modelling studies in the tropics that used a radiative-convective equilibrium setup (Bony et al., 2008; Risi et al., 2008; Blossey et al., 2010) produced a negative bias in heavy isotopes of precipitation (Bony et al., 2008; Torri et al., 2017).

2 Data and methods

2.1 Numerical simulations

We carried out three simulations with explicit convection and with different grid resolutions (10, 5, 1 km) using the regional model COSMO$_{iso}$. The implementation of isotopes in COSMO$_{iso}$ is achieved through a parallel water cycle for each of the two heavy isotopes (Pfahl et al., 2012), ^{1}H^{2}H^{16}O and ^{1}H$_2$^{18}O, which differs from the one of ^{1}H$_2$^{16}O by accounting for fractionation processes. The two additional water cycles are purely diagnostic and do not affect other model components. Fractionation processes during soil water evaporation are included through the coupling of COSMO$_{iso}$ and TERRA$_{iso}$, an isotope-enabled prognostic multilayer soil model (for details see Christner et al., 2018). Plant transpiration is treated as a non-fractionating process (Aemisegger et al., 2015). For evaporation from the ocean surface, the non-equilibrium fractionation factors from the smooth regime of Merlivat and Jouzel (1979) were used. Cloud processes were simulated by a one-moment microphysics scheme with a fixed number of cloud droplets ($N_c = 5 \times 10^8$ m^{-3}; Doms et al., 2011). The convection schemes of the model (Tiedtke, 1989; Theunert and Seifert, 2006) were disabled, meaning that deep and shallow convection were treated explicitly at the grid scale. A 20 s model time step was used, and hourly output was generated. A nested approach with spectral nudging of horizontal winds in the free troposphere was chosen to allow for a direct comparison to observations and to correctly capture influences from large-scale advection via the lateral boundary conditions.

The initial and lateral boundary conditions for the COSMO$_{iso}$ simulation with the coarsest spatial resolution, referred to as COSMO$_{iso, 10km}$, were taken at 6-hourly intervals from a simulation performed with the global model ECHAM6-wiso

(Cauquoin et al., 2019; Cauquoin and Werner, 2021). To reproduce the large-scale meteorological conditions during EUREC[4]A, ECHAM6-wiso was nudged (incl. surface pressure, temperature, and horizontal winds) towards ERA5 (Hersbach et al., 2020). The ECHAM6-wiso simulation was performed with parameterised convection and used the wind-dependent non-equilibrium fractionation factors from Merlivat and Jouzel (1979) for evaporation from the ocean surface. The global model was run with the spectral grid T127 (approx. $0.9°$ horizontal resolution) and 95 vertical levels.

COSMO$_{iso, 10km}$ has a horizontal resolution of $0.1°$, 40 vertical levels, and covers most of the North Atlantic (Fig. 2, Table 1). Horizontal winds above $850\,hPa$ in COSMO$_{iso,10km}$ were nudged towards ECHAM6-wiso. The spectral nudging technique ensured that the regional simulations remain close to the large-scale flow in the global model (Von Storch et al., 2000; Schubert-Frisius et al., 2017). The nudging included only zonal and meridional wave numbers of less than five. COSMO$_{iso,10km}$ covers the period from 6 January to 13 February 2020. The first 10 days are treated as spin up and are not included in the analysis.

The second COSMO$_{iso}$ simulation, COSMO$_{iso,5km}$, has a horizontal resolution of $0.05°$, 60 vertical levels, and covers a subset of the western North Atlantic, including the northern part of the South American continent (Fig. 2, Table 1). Initial and lateral boundary conditions were taken from COSMO$_{iso,10km}$ at hourly time steps. The spectral nudging technique was identical to the first COSMO$_{iso}$ simulation but nudged towards COSMO$_{iso,10km}$ horizontal winds above $850\,hPa$ instead of ECHAM6-wiso (using the same wave numbers as for the COSMO$_{iso,10km}$ albeit resulting in smaller wavelengths given the smaller model domain). COSMO$_{iso,5km}$ covers the period from 20 January to 13 February 2020. All simulated days are included in the analysis.

The third COSMO$_{iso}$ simulation, COSMO$_{iso,1km}$, has a horizontal resolution of $0.01°$, 60 vertical levels, and covers the focus area of EUREC[4]A's field activity (Fig. 2, Table 1). Initial and lateral boundary conditions were taken from COSMO$_{iso,5km}$ at hourly time steps. The spectral nudging, in this case, was directed towards COSMO$_{iso,5km}$ wind data. COSMO$_{iso,1km}$ covers the period from 20 January to 13 February 2020, and all simulated days are taken into account for the analysis.

2.2 Observational datasets

2.2.1 Satellite observations

The true colour images from the Visible Infrared Imaging Radiometer Suites (VIIRS) onboard the polar-orbiting satellite Suomi-NPP (National Polar-orbiting Partnership; NOAA, 2017) are used to assess the spatial organisation of clouds. The images were retrieved from the Worldview Snapshot application on NASA's Earthdata platform. For the same purpose, the data (with a spatial resolution of $500\,m$) from the visible channel of the satellite GOES-16, which is part of NOAA and NASA's Geostationary Operational Environmental Satellite (GOES) – R Series, were downloaded from the EUREC[4]A data catalogue.

The Integrated Multi-satellitE Retrievals (IMERG) from the Global Precipitation Measurement (GPM) mission is used for the evaluation of the spatial distribution of precipitation. We refer to the data, which were obtained from NASA's Earthdata platform, simply as GPM. Here, the half-hourly (Huffman et al., 2019a) and daily (Huffman et al., 2019b) precipitation estimates are used.

2.2.2 Aircraft observations

Several observational datasets that were created during EUREC[4]A onboard the ATR and the German aircraft HALO (High Altitude and Long Range Research Aircraft; Konow et al., 2021) are used. All of them are available on AERIS (https://eurec4a.aeris-data.fr, last access: 18 September 2022). The ATR conducted 19 flights on 11 days from 25 January to 13 February 2020. A flight typically lasted four to five hours and consisted of repetitive flight patterns. They included the ferry from the Grantley Adams International Airport of Barbados into the western half of the so-called EUREC[4]A circle (see definition below), followed by two or three rectangles ($120\,\mathrm{km} \times 20\,\mathrm{km}$) at cloud base, one or two L-patterns flown either near the top or the middle of the subcloud layer, occasionally a surface leg at an altitude of about $60\,\mathrm{m}$, and the ferry back to the airport, which often included a short leg above the trade inversion to sample free tropospheric air. The ATR flights were closely coordinated with those of the HALO, which performed 15 flights from 19 January to 18 February 2020. Each HALO flight lasted about nine hours, seven of which the aircraft spent at upper altitudes ($\sim 10\,\mathrm{km}$) in the EUREC[4]A circle centred at $13.3°\,\mathrm{N}$, $57.7°\,\mathrm{W}$ (roughly $150\,\mathrm{km}$ east of Barbados), with a diameter of $\sim 220\,\mathrm{km}$. When the HALO was flying this circle, dropsondes were launched frequently to estimate the large-scale vertical motions (Bony and Stevens, 2019; Stevens et al., 2021; Konow et al., 2021; George et al., 2021a, 2023). The HALO dropsonde data (George et al., 2021b) is used for the evaluation of the COSMO$_{\mathrm{iso}}$ simulations.

The ATR's position and the corresponding standard meteorological variables (e.g., pressure, temperature, wind components) are reported in the CORE dataset (CNRM/TRAMM et al., 2021; Bony et al., 2022, 1 Hz temporal resolution). Considering that the ATR moved roughly $100\,\mathrm{m\,s^{-1}}$, the spatial resolution of its observations amounts to $100\,\mathrm{m}$ for the 1 Hz data and to $1\,\mathrm{km}$ for the 10 s-averaged data.

The ATR-based BASTALIAS dataset (Delanoë et al., 2021) is used for the evaluation of the cloud fraction at cloud base. It is a data product resulting from the combined observations of the Doppler cloud radar BASTA (Bistatic Radar System for Atmospheric Studies; Delanoë et al., 2016) and the lightweight backscatter lidar ALiAS (Airborne Lidar for Atmospheric Studies; Chazette et al., 2020) that were horizontally probing the surroundings of the ATR.

The ATR-based PMA dataset (Microphysics Airborne Platform; Coutris, 2021, 1 Hz temporal resolution) is used for the microphysical characterisation of the clouds and their surroundings. It contains the combined observations from an optical-scattering droplet spectrometer (Lance et al., 2010) and an optical array stereo probe imager (Lawson et al., 2006), together sizing and counting the droplets with diameters from $2\,\mathrm{\mu m}$ to $2\,\mathrm{mm}$. From the size distribution composites of the two instruments, the total concentration of particles, the median volume diameter, the liquid water content (LWC), and flags indicating the presence of cloud, drizzle or raindrops are inferred. Cloud drops are identified if the LWC exceeds $10\,\mathrm{mg\,m^{-3}}$ and the particle diameter remains below $100\,\mathrm{\mu m}$; drizzle if the LWC exceeds $10\,\mathrm{mg\,m^{-3}}$ and the particle diameter is at least $100\,\mathrm{\mu m}$ but remains below $500\,\mathrm{\mu m}$; and rain if the LWC exceeds $10\,\mathrm{mg\,m^{-3}}$ and the particle diameter exceeds $500\,\mathrm{\mu m}$.

Finally, the ATR-based Picarro dataset (Aemisegger et al., 2021a) is used for the humidity conditions and the isotopic composition of the vapour, which were both measured with a customised cavity ring-down spectrometer L2130-*i* from the manufacturer Picarro with a sampling frequency of 1 Hz (more technical details in Bailey et al., 2023). The Picarro dataset

provides a flag indicating data points of questionable quality (e.g., due to inlet wetting). These data points were removed for the analyses shown in this paper. Moreover, only 10 s-averages (for a better noise-to-signal ratio) are used, which corresponds

285 to an effective horizontal resolution of 1 km and thus makes the observations comparable to the COSMO$_{iso,1km}$ simulation.

Note that the ATR observations have been thoroughly evaluated in Bony et al. (2022). They found that the different ATR measurements of humidity, wind and cloud fraction are in good agreement with each other and with HALO dropsondes.

2.3 Definition of the cloud base level

The ATR data comes with an official flight segmentation (Bony et al., 2022). Here, we use the data that is labelled as cloud

290 base. The targeted cloud base level was determined before each flight based on ceilometer, radio sounding, and surface weather observations as well as satellite imagery. During the flight, the altitude was fine-tuned based on real-time measurements and visual impressions. To ensure that only the data points collected at the base of shallow clouds are taken into account, we apply the criterion that only data points with altitudes below 1.3 km (threshold for the distinction between shallow and deep clouds used in Vial et al., 2019) are treated as cloud base observations. The resulting cloud base mean and median pressure for each

295 flight is shown in Fig. 3.

For comparison with the ATR data, a model level representing cloud base in the domain 54.5-61° W and 11-16° N is defined for the COSMO$_{iso}$ simulations. This model level is identified for every hourly time step individually, following the steps below:

1. Every vertical profile in the domain 54.5-61° W and 11-16° N is checked for cloud water content exceeding $10 \, \mathrm{mg \, kg^{-1}}$ (threshold for the detection of clouds used in Vial et al., 2019) at any model level below 1.3 km. The lowest model level

300 meeting this criterion is taken as the cloud base of the respective vertical profile. If the profile does not contain any clouds, it is ignored in the subsequent step.

2. In order to extract cloud base conditions, we determine one cloud base model level for the domain 54.5-61° W and 11-16° N by calculating the median over the cloud base model levels identified for cloudy profiles in the previous step.

3. Steps 1 and 2 are repeated for every hourly time step of the simulated time period. The resulting time series of cloud

305 base model levels can then be used to extract cloud base variables (as, e.g., pressure shown in Fig. 3) from the COSMO$_{iso}$ simulations.

The cloud base levels identified in the three COSMO$_{iso}$ simulations, using the procedure described above, alternate between the model levels 29 (783 m) and 30 (970 m) for COSMO$_{iso,10km}$ and 47 (761 m), 48 (914 m), and 49 (1082 m) for COSMO$_{iso,5km}$ and COSMO$_{iso,1km}$. COSMO$_{iso}$ underestimates the flight-to-flight variability of the cloud base altitude in comparison to the

310 ATR data (Fig. 3). Nevertheless, the identified COSMO$_{iso}$ cloud base altitudes are in the range of the values observed by the ATR during the cloud base flight segments (Fig. 3b). Therefore, the COSMO$_{iso}$ cloud base data is assumed to be comparable to the ATR cloud base data.

3 Model evaluation and spatial resolution effects

EUREC[4]A offers a unique opportunity to perform a thorough evaluation of model simulations targeted at shallow trade-wind cumulus clouds in a process-oriented way thanks to the wealth of available observational data. In the following, variables from the different $COSMO_{iso}$ simulations, which are considered to be relevant for the formation of shallow cumulus clouds in the trade-wind region, are selected and compared to observations. Simultaneously, the effect of the spatial model resolution is assessed, without upscaling outputs from the higher resolution simulations to the lowest resolution. All comparisons are performed with the data in the selected evaluation domain 54.5-61° W and 11-16° N (corresponding approximatively the domain of the $COSMO_{iso,1km}$ simulation). This domain was chosen large enough to provide a robust statistical basis to analyse cloud properties while remaining small enough to contain a fairly homogeneous cloud field in terms of cloud patterns.

3.1 Clouds and precipitation

The $COSMO_{iso}$ simulations are all able to distinguish between different mesoscale organisation patterns (Fig. 4; other dates can be found in Supplement 1). Note that the model is not expected to produce clouds identical to the ones in the observations, but it is plausible that the simulations reproduce the mesoscale cloud organisation patterns since they are linked to distinct large-scale environmental conditions (Bony et al., 2020). The elongated Fish clouds on 22 and 26 January 2020 are reproduced by $COSMO_{iso}$, although, at a different location than depicted by the satellite image, hinting towards slight errors in the propagation speed of the banded clouds associated with trailing cold fronts (see Villiger et al., 2022, for the link between Fish clouds and trailing cold fronts). On 12 February 2020, the satellite shows large Flower clouds, which are not clearly identifiable in the $COSMO_{iso}$ output. Nevertheless, if viewed over a larger domain (Supplement 1), aggregated cloud liquid water content objects are visible, resembling Flower clouds. Smaller Flower clouds, as on 2 February 2020, are clearly reproduced by the model. Finally, more finely structured cloud patterns recognised as Gravel and Sugar, as they appear on 8 and 9 February 2020 in the satellite observations, are also reproduced by $COSMO_{iso}$. Note that the cloud classification used above largely agrees with the patterns identified by Schulz (2022, their Fig. 7a).

Overall, the spatial cloud patterns from the satellite images are qualitatively better reproduced by the simulations if the horizontal model resolution is increased. This holds especially for the finely structured clouds on 8 and 9 February 2020. For these cases, only $COSMO_{iso,1km}$ shows the Gravel-typical cold pool activity in the form of cloud liquid water arranged in arc-like structures around clear-sky patches. An exception is 12 February 2020, a day on which the organisation of the clouds in large Flowers is partially lost with the higher resolution of $COSMO_{iso,1km}$.

The four mesoscale organisation patterns discussed above are known to be characterised by different low-level cloud fractions (Bony et al., 2020), peaking at different times of the day (Vial et al., 2021). Therefore, the temporal evolution of cloud fraction at cloud base is another measure to evaluate the performance of the three $COSMO_{iso}$ simulations. In all three simulations, the variable is characterised by a pronounced diel cycle, with maximum cloud fraction during local early morning (Fig. 5; agreeing with Vial et al., 2019). The magnitude of the cloud fraction, however, is very different for the three simulations. The cloud fraction is largest in $COSMO_{iso,10km}$ and becomes progressively smaller with higher resolution until reaching minimum

values in $COSMO_{iso,1km}$. This behaviour is expected for clouds that are smaller than the grid resolution (explanatory sketch in Supplement 1) and has been observed in earlier studies (Hentgen, 2019; Heim et al., 2021). The motivation for displaying cloud fractions from all three simulations is to assess at which resolution we obtain cloud fractions that approximate the observations within their uncertainty range. The comparison to the ATR-based cloud fraction estimates (symbols in Fig. 5) shows that

350 $COSMO_{iso,10km}$ and (for most of the flights) $COSMO_{iso,5km}$ overestimate cloud fraction. Contrastingly, $COSMO_{iso,1km}$ agrees well with $ATR_{BASTALIAS,cloud+drizzle}$ except for the first flight on 2 February and the two flights on 7 February 2020. On these two dates, however, coarse cloud patterns (Flower and Fish; not shown) were present, meaning that the ATR likely underestimated cloud fraction due to its limited sampling area.

In the model, liquid water droplets in a grid cell are separated into the non-sedimenting cloud water content and the sedi-

355 menting rain water content, respectively. For comparison to the ATR-based liquid water content (LWC) measurements provided in the PMA dataset (Fig. 6a), the simulated cloud and rain water contents are summed up. Consistent with cloud base cloud fraction, the fractions of data points with the $LWC > 10\,mg\,kg^{-1}$ are decreasing with increasing spatial model resolution and overall higher in the simulations than in the observations. $COSMO_{iso,10km}$ meets the threshold at $22.3\,\%$ of the data points, $COSMO_{iso,5km}$ at $14.0\,\%$, $COSMO_{iso,1km}$ at $8.3\,\%$, and the PMA at $5.6\,\%$.

360 The LWC at cloud base is generally overestimated in $COSMO_{iso}$ compared to the observations, but the observed values are approached with increasing horizontal model resolution (Fig. 6a). Possibly the number of cloud droplets (N_c), which is fixed to $5 \times 10^8\,m^{-3}$ in the one-moment cloud scheme (Sect. 2), contributes to the offset between simulations and observations. This value is located at the upper end of the observed droplet number distribution (Fig. 6b). A high N_c leads to a reduction of the cloud droplets' size since the available liquid water is distributed over more droplets. This, in turn, reduces the efficiency of

365 the collisional growth process and can delay or even prevent the formation of precipitation (see, e.g., Eirund et al., 2022). In other words, N_c directly affects the autoconversion rate, the rate at which $COSMO_{iso}$ turns cloud water into rain water. Thus, if the parameter were adjusted towards the observations, precipitation would likely be formed more readily (efficiently removing liquid water from the atmosphere), possibly affecting below-cloud evaporation of hydrometeors and surface precipitation. It remains to be tested how $COSMO_{iso}$ would react to a change of N_c in the considered regime.

370 Figure 7 shows the rain water that ultimately reaches the surface. Compared to the satellite product, precipitation falls over smaller areas in $COSMO_{iso}$, even if the resolutions of the compared datasets are comparable ($COSMO_{iso,10km}$ vs. GPM; keep in mind that the satellite product is an estimate of surface precipitation and should not be taken as the absolute truth). The patches with intense rainfall become smaller and more frequent with finer model resolution. Most likely, this is due to an intensification of updrafts with finer model resolution.

375 Considering that the spatial distribution of precipitation in $COSMO_{iso}$ strongly contrasts with the satellite product (Fig. 7), the rainfall amounts over the evaluation domain are surprisingly close to each other (Fig. 8). It depends on the dominant cloud pattern, whether $COSMO_{iso}$ is over-/underestimating precipitation and whether the grid resolution has a systematic effect. From 20 to 24 January, when Fish clouds were dominant (Schulz, 2022, their Fig. 7), the simulations show a distinct diel cycle of precipitation (peaking shortly after the peak in cloud fraction; agreeing with Vial et al., 2016), while the satellite reports a period

380 of continuous rainfall (Fig. 8a). Over a $24\,h$ window (Fig. 8b), this leads to a rainfall underestimation on 22 and 24 January

2020 by all three simulations, while on 20, 21, and 23 January 2020 alternately one of the three simulations best matches with the satellite. The exaggerated diel cycle of precipitation in the simulations highlights potentially too strong radiatively driven mesoscale circulations that may slow down the propagation of the trailing cold front into the tropics (cf., shifted position of the Fish cloud in the simulations compared to the observations; Fig. 4). Note that this period was characterised by anomalous large-scale transport (Villiger et al., 2022) with the precipitation-triggering process being dominated by large-scale convergence instead of isolated cells of shallow convection.

From 25 January to 5 February 2020, when the clouds either are classified as Sugar or remain unclassified (Schulz, 2022, their Fig. 7), all COSMO$_{iso}$ simulations overestimate the rainfall amount. The overestimation is smaller the finer the spatial model resolutions. From 6 February 2020 onwards, the rainfall amount is increased in all four datasets, but again none of the three simulations is exceptionally accurate. Much more, it depends on the date, which of the simulations is closest to the observations. This final period of the campaign was again characterised by anomalous large-scale transport (weak extratropical dry intrusion on 7 and 8 February 2020 and tropical mid-level detrainment starting on 12 February 2020 discussed in Villiger et al., 2022). The dominant cloud patterns in the considered domain were Fish clouds (7 to 8 February 2020) followed by Flower clouds (10 to 12 February 2020; Schulz, 2022, their Fig. 7). Besides 20 to 26 January 2020, the diel cycle of precipitation from COSMO$_{iso}$ agrees with the one from the satellite observations (Fig. 8a).

The variables considered up to this point show a high dependence on the horizontal model resolution. This is not surprising since the three COSMO$_{iso}$ simulations are in the so-called "grey zone" of resolutions, where only some of the scales involved in convective motions are resolved (Wyngaard, 2004), while the sub-grid-scale processes are taken care of by the turbulence scheme. This interplay between convection and turbulence leads to a strong resolution-dependency of variables associated with convection (Hanley et al., 2015; Jeevanjee, 2017). Whether the increase in resolution leads to a better agreement with observations or not depends on the variable. For instance, the representation of the cloud parameters in the simulations is improved with finer resolution, while this is not the case for the spatial distribution of precipitation. For evaluations of Large-Eddy simulations with EUREC[4]A observations at even finer grid spacings than our COSMO$_{iso}$ simulations, we refer to Schulz and Stevens (2023).

3.2 Water vapour isotopes, humidity and temperature

In contrast to the highly skewed variables discussed in the previous section, the close-to-normally distributed variables considered in this section are largely indifferent to model resolution (Fig. 9 and 10). This limited resolution dependence results from the fact that here we consider the full distribution of these variables and do not focus on the properties of subgrid-scale features as in the previous section.

The (close to) saturated layer extends over ~950-900 hPa in COSMO$_{iso}$ and over ~950-820 hPa in the measurements from the downward profiles of the ATR, the measurements from the HALO dropsondes, the ERA5 reanalysis, and the ECHAM6-wiso simulation (Fig. 9e). Using high relative humidity ($>80\,\%$) as a proxy for the presence of clouds (since observation-based cloud fraction estimates are only available at cloud base), we conclude that COSMO$_{iso}$ generally has a shallower cloud layer than the other datasets and presumably has difficulties in producing cloud top anvils accurately. This behaviour is also observed

415 in other regional model simulations of shallow trade-wind clouds (Beucher et al., 2022, note that they included a shallow convection parametrisation, while we do not). In our COSMO$_{iso}$ simulations, the rather coarse vertical resolution (Sect. 2) might be a reason for the too shallow representation of clouds.

 Compared to the measurements from the downward profiles of the ATR, COSMO$_{iso}$ has a clear cold-dry bias of 1.8-2.1° C and 1.6-1.7 g kg^{-1} (Fig. 9a,d and Table 2) in the lower troposphere (from the surface to ∼850 hPa). The HALO dropsonde

420 measurements, ERA5, and ECHAM6-wiso (in which the temperature was nudged) match better with the ATR measurements. ECHAM6-wiso even has a weak moist bias in the cloud layer from about 940 hPa to 700 hPa, meaning that COSMO$_{iso}$ produces its dry-cold bias independently of the ECHAM6-wiso boundary data. Furthermore, COSMO$_{iso}$'s temperature profile shows a slightly deeper adiabatic layer (up to 925 hPa; Fig. 9b,c), which suggests a deeper subcloud layer than in the observations, ERA5, or ECHAM6-wiso (cf., Beucher et al., 2022, who found the same in their simulations). The deeper adiabatic layer

425 leads to a more stable cloud layer, with a stronger increase in virtual potential temperature with height (Fig. 9b) in COSMO$_{iso}$ than in the other datasets. This matches the finding above that COSMO$_{iso}$ has fewer deep clouds (reaching above 850 hPa).

 The cold-dry bias in COSMO$_{iso}$ is associated with slightly too depleted vapour compared to the ATR measurements (Fig. 9f,g and Table 2). In contrast, COSMO$_{iso}$ produces too enriched vapour in the free troposphere (above ∼800 hPa). ECHAM6-wiso shows slightly too enriched vapour in the cloud layer and a more accurate decrease in heavy isotopes above cloud tops compared

430 to the ATR observations than COSMO$_{iso}$. The d-excess in vapour (Fig. 9h) is almost constant throughout the vertical column in ECHAM6-wiso and COSMO$_{iso}$. While the values in ECHAM6-wiso overlap with the ATR data (at least up to ∼750 hPa), COSMO$_{iso}$ is shifted towards higher values by ∼4.5‰ (Fig. 9h; the bias is visible up to 200 hPa, not shown). Given the fact that ECHAM6-wiso does not have the same isotope biases as COSMO$_{iso}$, it follows that COSMO$_{iso}$ must produce them within the regional model domain.

435 The low-level biases in temperature, humidity, and δ^2H are physically consistent. Due to the cold bias alone, a dry and a depletion bias can be expected. In fact, about two-thirds of the bias of δ^2H and one-fourth of the bias of δ^{18}O can be explained by the temperature bias (via the temperature-dependency of equilibrium fractionation). This also means that isotopes can only provide limited information about the origin of the low-level biases. Different mechanisms possibly cause some of the identified biases at low levels: (1) An overestimation of the evaporation of hydrometeors would explain a low-level cold

440 bias (Fig. 9a). This, however, would simultaneously lead to a moistening, which is not consistent with the dry bias in the COSMO$_{iso}$ profiles (Fig. 9d). (2) A too strong convective mixing (across the full depth of the shown profiles), which transports dry-depleted air downward and moist-enriched air upwards, would fit the observed low-level dry-depletion bias, but only partly the conditions in the free troposphere with the too enriched vapour but no moist bias (Fig. 9d,f,g). Moreover, such a mixing process is expected to lead to a low-level warm bias (downward transport of high potential temperature and upward transport of low potential temperature) and not a low-level cold bias, as observed here (Fig. 9a). Lastly, the underestimation of clouds

445 reaching above 900 hPa (Fig. 9e) does not directly support too strong mixing. If convective mixing was too strong, the few higher-reaching clouds in COSMO$_{iso}$ would need to mix much more efficiently than their more abundant shallow counterparts in the other datasets. (3) Another possible explanation for the cold bias that has not been investigated here but has been found in earlier COSMO evaluation studies (Heim et al., 2021) is too high outgoing longwave radiation (top-of-the-atmosphere) due to

a too dry free troposphere or too few low-level clouds. However, we don't observe a dry bias in the free troposphere (Fig. 9d). Furthermore, COSMO$_{iso}$'s cloud base cloud fractions are rather too high compared to the ATR-based estimates (Fig. 5). (4) Finally, we cannot rule out that the biases are caused by the spectral nudging (horizontal winds above 850 hPa) applied for the simulations (e.g., Wehrli et al., 2018; Sun et al., 2019). However, in another study (Beucher et al., 2022), a cold-dry bias was also found in free-running simulations that were performed without spectral nudging.

Two mechanisms might explain the enrichment bias above 800 hPa (Fig. 9f,g): (1) Too small conversion efficiencies (i.e., too low autoconversion rate within the cloud column) could explain an enrichment bias of about 10 ‰ in δ^2H in the cloud top outflow region of shallow cumulus clouds (see Rayleigh model including precipitation efficiency in Noone, 2012). However, the magnitude of the bias observed here (offset of roughly 60 ‰ for δ^2H) is too large to solely be explained by this effect. (2) Another possible explanation is that too strong cloud top evaporation occurs due to too strong turbulent mixing at the trade inversion. This could lead to only weak moistening but strong enrichment due to close to total (thus non-fractionating) droplet evaporation. The large isotope bias collocated with a very small or inexistent humidity bias in the region of the trade inversion and the level of the deepest cloud tops illustrates that the isotopes contain additional, complementary information for assessing some model biases.

Possible reasons for a d-excess bias (Fig. 9h) are wrong near-surface humidity gradients or the use of inappropriate non-equilibrium fractionation factors. A comparison to ship-based observations from EUREC[4]A (not shown) revealed that in our case, too low relative humidity with respect to sea surface temperatures accounts for a large part of the identified d-excess bias. Since the sea surface temperatures in COSMO$_{iso}$ are identical to the ones in ECHAM6-wiso, the too low relative humidity must result from the dry bias discussed above. The remaining share of the d-excess bias might be due to the choice of the non-equilibrium fractionation factor in the model simulations. Evidence is available that the chosen value in COSMO$_{iso}$ can still be improved compared to observations (Zannoni et al., 2022). The wind-dependent formulation of Merlivat and Jouzel (1979), used in ECHAM6-wiso, and the output from a sensitivity experiment with COSMO$_{iso}$ with a new wind-dependent formulation of the non-equilibrium fractionation factor from work done by Dütsch (2021, see experimental simulation shown in Supplement 4) yield d-excess values closer to the ATR observations than the evaluated COSMO$_{iso}$ simulations.

The slightly deeper subcloud layer (adiabatically stratified; Fig. 9b,c) of COSMO$_{iso}$ in comparison to the ATR and HALO observations, as well as ECHAM6-wiso and ERA5 data, is likely due to too strong turbulent mixing in the subcloud layer and at cloud base. Additionally, the differences in the vertical resolution of the compared datasets could play a role in the observed differences in the profiles. Indeed, ERA5 has roughly twice as many model levels in the layer 1000 to 850 hPa as the three COSMO$_{iso}$ simulations. However, ECHAM6-wiso has fewer levels in the considered layer. Thus, it is not plausible that the deeper adiabatic layer in COSMO$_{iso}$ originates only from a vertical resolution effect. The above-stated hypothesis of a too strong turbulent mixing at low levels in the simulations is confirmed by smaller vertical differences between cloud base and subcloud values of humidity, δ^2H, and d-excess in COSMO$_{iso}$ compared to the ATR as shown in Villiger (2022).

After having discussed various biases, we want to evaluate the development of isotope variables over time. To this end, we focus on cloud base and compare the flight-to-flight variability found in the ATR data to the simulations. High correlations (Table 3) between the two datasets (Fig. 10) suggest that the observed temporal evolution of δ^2H and d-excess is well captured

by the three COSMO$_{iso}$ simulations. Another similarity between observations and simulations is that δ^2H (Fig. 10 a) generally shows a larger in-flight spatial variability at cloud base compared to the flight-to-flight temporal (synoptic) variability. This highlights the importance of mesoscale variability at cloud base, which is the topic of the following Sect. 4.

4 Drivers of the horizontal variability of isotopes at cloud base

The detailed examination of the observed and simulated isotope signals at cloud base in the previous section (Fig. 10) showed that δ^2H has a large spatial (in-flight) and temporal (flight-to-flight) variability, while the d-excess varies mainly from flight to flight. The flight-to-flight variability is related to the variability of the synoptic circulation. From earlier studies, it is known that the d-excess is negatively correlated with the distance to the moisture source (Aemisegger et al., 2021b), which is closely linked to the prevailing large-scale circulation. However, the drivers behind the mesoscale (in-flight) variability of isotopes at cloud base have not been investigated so far. Here, this topic is addressed by looking at the isotopic characteristics of different environments at cloud base. First, features at the cloud base level are defined in the observational and simulated data based on a case study on 2 February 2020. Second, the whole datasets are used for a statistical description of these features. Based on the findings of Sect. 3, only the data from the COSMO$_{iso,1km}$ simulation is considered here. Compared to the lower-resolution COSMO$_{iso}$ simulations, COSMO$_{iso,1km}$ reproduces the spatial cloud patterns, cloud base cloud fraction and liquid water content better.

4.1 Detailed case study

On 2 February 2020, the ATR conducted research flight 9 (RF09) and 10 (RF10) (Bony et al., 2022). Here, we focus on RF10 for an illustrative case study. RF10 included three rectangles at cloud base. During these, the aircraft sampled clear-sky environments at both edges of the rectangle and a Flower cloud in-between (Fig. 11a,b and 12e). Beneath the Flower cloud, a cold pool is spreading radially away from the cloud centre (Fig. 11a,b). The edges of the cold pool are clearly marked by shallow clouds arranged in an arc at the edge of a clear-sky patch visible in the satellite image (blue dashed lines in Fig. 11a,b). The ATR cloud base rectangle is positioned such that its southern tip lies inside (or above) the cold pool, while its northern tip lies outside.

The four sondes dropped from the HALO in proximity to the flight track of the ATR demonstrate how different the environments in and around the Flower cloud were (note that the snapshots of the satellite images in Fig. 11 do not correspond to the exact launch times of the dropsondes, which require approximately 10 min to reach the surface assuming a launch altitude of 10 km and an average fall speed of 16 m s^{-1} based on the value range suggested in George et al., 2021a). Looking at the measured profiles from the perspective of the ATR flying at cloud base ($\sim$940 hPa; sandy line in Fig. 12e), one can distinguish between a profile with dry-warm conditions at cloud base (red profile in Fig. 11), two profiles with humid-cold conditions at cloud base (yellow and green profiles in Fig. 11), and a profile with a deep cloud layer aloft (blue profile in Fig. 11).

The blue profile with the deep cloud layer aloft, near the centre of the Flower cloud, bears the imprint of the cold pool in the subcloud layer. It is characterized by a saturated layer from 900 to 700 hPa (Fig. 11c), which most likely led to the

formation of rain (also observed by the ATR, see Fig. 12e). Below the cloud layer, isothermal conditions prevail (Fig. 11e) while relative humidity (Fig. 11c-d) rapidly decreases (promoting the evaporation of falling hydrometeors). At about 975 hPa, relative humidity reaches a local minimum value (Fig. 11c), which is co-located with a temperature drop of about 2 K (Fig. 11e),
520 pointing towards a strong evaporatively triggered cold pool. A similar, although weaker, temperature drop is visible in the green profile, located at the edge of the cold pool. The comparably shallow mixed-layer height in the two (blue and green) soundings matches the definition of cold pool soundings applied in Touzé-Peiffer et al. (2022).

Although the yellow profile is located inside the cold pool (Fig. 11a,b), it doesn't show the cold-pool characteristic temperature drop near the surface. Possibly, the differences between the green and yellow profiles are related to their position
525 relative to the cold pool centre. Towards the east, where the yellow profile is located, the radially spreading cold pool moves against the general flow of the easterlies, seemingly leading to a quick alteration of the air's cold-pool characteristics. Lastly, as expected from its position outside the cold pool (Fig. 11a,b), the red profile is totally unaffected by the cold air spreading near the surface. Instead, it shows an adiabatic profile, typical for the well-mixed subcloud layer.

At the altitude flown by the ATR, the largest horizontal contrasts in humidity and temperature are found between the cloudy
530 (green) and warm clear-sky (red) dropsonde profiles, which match with the observations of the ATR (Fig. 12). In the clear-sky background environments at the northern edge of the rectangle (at 13.8° N; Fig. 12f), specific humidity and δ^2H in vapour reach minimum values (Fig. 12b), while a level-relative maximum is reached in temperature (Fig. 12c). On the long side of the rectangle, inside cloudy flight sections (Fig. 12e and grey shadings in all panels), specific humidity and δ^2H in vapour are higher, and temperature is lower compared to their respective mean values. Regarding the turbulence intensity indicated
535 by the vertical velocity fluctuations (relative to the flight-segment mean; Fig. 12d), the cloudy sections are quite turbulent with amplitudes of fluctuations up to $2 \, \text{m s}^{-1}$, while the clear-sky background environments exhibit weak turbulence close to laminar flow conditions. For the d-excess (Fig. 12a), it is not possible to make a statement about a possible influence of the immediate surroundings (cloud vs. clear-sky) due to the noise of the measurements.

The fact that the clear-sky environment north of the Flower cloud distinguishes itself from the one in the south by lower
540 humidity and higher potential temperature leads to the conclusion that it is a region where air masses subside (see the location of the ATR icon in the schematic Fig. 1a), balancing the upward air mass transport inside the clouds. The clear-sky environment at the southern edge of the cloud base rectangle has no particular anomalies in any observed variables. Possibly the characteristics of this southern clear-sky environment are shaped by dissipated former clouds (i.e., high humidity; cf., Albright et al., 2022).

Coming back to the original purpose of this section (i.e., to study the in-flight isotope variability at cloud base), the cloud
545 base time series of the ATR (Fig. 12) clearly show that the variability in δ^2H in the vapour is driven by the contrast between cloudy and dry-warm, clear-sky background environments. Thus, the question arises whether these patterns are reproduced by the COSMO$_{\text{iso}}$ simulations.

On 2 February 2020 at 15 UTC, a large, partially precipitating cloud, resembling a Flower cloud, appears in COSMO$_{\text{iso,1km}}$ just north of the ATR flight track (Fig. 13). The humid environment of the cloud is flanked to the west and east by dry-warm
550 patches (Fig. 13a,b and 14a). These dry-warm blobs are characterised by low δ^2H (Fig. 13c and 14b) and low d-excess (Fig. 13d and 14c) in vapour. Thus, COSMO$_{\text{iso,1km}}$ produces dry-warm features at cloud base, which match the above description based

on the ATR and HALO dropsonde observations. Furthermore, it is the contrast between the dry-warm blob and the cloudy environments (Fig. 13c) that shape the width of the distribution of simulated δ^2H values, similar to what has been observed in the ATR cloud base measurements.

555   Putting these findings into the context of the theory summarised in the introduction (Sect. 1), it follows that the variability of δ^2H at cloud base mirrors the mesoscale transport as cloudy environments are more enriched (updraft of freshly evaporated moisture) compared to dry-warm environments (subsidence of depleted vapour from aloft). This, however, does not mean that microphysical processes do not influence δ^2H as well. Depletion of the vapour inside the cloud due to the formation of liquid droplets most likely reduces the observed horizontal contrasts between cloudy and dry-warm environments at cloud base. The

560 imprint of these local processes on δ^2H in the vapour is, however, much smaller than the signal of the mesoscale transport that reflects the cloud processing integrated over the mesoscale moisture cycle.

4.2   Statistical evaluation

To assess the robustness of the insights gained from the 2 February 2020 case, objective criteria to detect the different cloud base environments in the two datasets are introduced. For the ATR, the flags from the PMA dataset (Bony et al., 2022) are used

565 to distinguish between time steps in clear-sky conditions and time steps with liquid water droplets in the surroundings. The following categories are defined (see the percentage of data points in each category in Table 4):

- *clear*: time steps with no PMA flags,

- *cloud*: time steps with the PMA flags indicating clouds, but neither rain nor drizzle,

- *cloud-rain*: time steps with the PMA flags indicating clouds and rain.

570 Note that the PMA flags indicate rain but no clouds for $1.3\,\%$ of the time (compared to $0.2\,\%$ indicating rain and clouds). These time steps are excluded from the categorisation because they complicate the interpretation of the isotope signal (highly depending on the formation altitude of the rain). The time steps identified as *clear* are relabelled as *dry-warm* if they have a negative anomaly in specific humidity (q) and a positive anomaly in potential temperature (θ). The anomalies are defined relative to each flight's mean $\overline{X}$ and standard deviation $\sigma(X)$ of the respective variable X at cloud base with the criteria

575 $q < \overline{q} - \sigma(q)$ and $\theta > \overline{\theta} + \sigma(\theta)$ (e.g., red lines in Fig. 12b,c), which must be fulfilled for a time step to be assigned to the *dry-warm* category. The two categories *clear* and *dry-warm* are exclusive (i.e., data points identified as *dry-warm* are removed from the *clear* category).

    For the COSMO$_{\text{iso,1km}}$ data, the cloud base grid points in the evaluation domain 54.5-61° W and 11-16° N are stratified according to their cloud and rain water contents:

580 - *clear*: grid points with cloud water content below $10\,\text{mg}\,\text{kg}^{-1}$ and rain water content below $1\,\text{mg}\,\text{kg}^{-1}$,

- *cloud*: grid points with cloud water content exceeding $10\,\text{mg}\,\text{kg}^{-1}$ (black contours in Fig. 13) and rain water content below $1\,\text{mg}\,\text{kg}^{-1}$,

– *cloud-rain*: grid points with cloud water content exceeding $10\,\mathrm{mg\,kg^{-1}}$ (black contours in Fig. 13) and rain water content exceeding $1\,\mathrm{mg\,kg^{-1}}$ (green contours in Fig. 13).

585 From the *clear* category, the grid points with a negative anomaly in q and a positive anomaly in θ are reassigned to the *dry-warm* category, similarly as it is done for the measurements. Here, the anomalies are defined grid-point-wise, by removing the daily cycle. The hour-of-the-day mean and standard deviation for each grid point are calculated over the whole simulated period (20 January to 13 February 2020). Note that for the ATR observations, the effect of the daily cycle could not be subtracted because too few observations are available, in particular at night, preventing a robust estimation of the daily cycle. *Dry-warm*

590 grid points are then selected using the following criteria: $q_{i,t} < \overline{q_{i,h(t)}} - \sigma(q_{i,h(t)})$ and $\theta_{i,t} > \overline{\theta_{i,h(t)}} + \sigma(\theta_{i,h(t)})$ (i indicates the grid points in the evaluation domain, t the hourly time steps of the simulated period, and $h(t)$ the hour of the day corresponding to the time step t). An overview of the number of identified *dry-warm* grid points per time step is given in Supplement 3 and for the case study on 2 February 2020 the *dry-warm* grid points are shown as red dots in Fig. 13a. The grid points with rain water content exceeding $1\,\mathrm{mg\,kg^{-1}}$ but remaining below the cloud water content threshold of $10\,\mathrm{mg\,kg^{-1}}$ are not assigned to

595 any category. Such grid points make up $0.6\,\%$ of the cloud base grid points over the hourly time steps closest to the ATR flights (compared to $0.3\,\%$ with rain and clouds).

Applying the above-described categorisation to the full ATR and COSMO$_{\mathrm{iso,1km}}$ datasets yields similar gradients in the cloud base environments (Fig. 15) as observed on 2 February 2020 (Fig. 12, 13 and 14). Both datasets describe a gradient from dry-depleted to moist-enriched conditions, going from *dry-warm* to *cloud/cloud-rain* environments (Fig. 15a,b). The specific

600 humidity anomalies of the different cloud base environments are comparable in the ATR and the COSMO$_{\mathrm{iso}}$ data (Table 5). The δ^2H anomalies are comparable for the *dry-warm* category. The *cloud* and *cloud-rain* environments, however, are more enriched relative to the level mean in the ATR data than in the COSMO$_{\mathrm{iso}}$ data.

Due to the chosen statistical approach, we can now also identify contrasts in the d-excess at cloud base with negative anomalies in the *dry-warm* blobs and positive anomalies inside clouds (Fig. 15c). However, the d-excess distributions of the

605 different cloud base environments overlap a lot, both in the ATR data and the simulations (especially if the full width of the distribution is considered as in Supplement 3). The direction of the observed d-excess gradients does not fit the general expectations of high d-excess being associated with low relative humidity and vice versa (Bony et al., 2008). Looking at the vertical sections on 2 February 2020 (Fig. 14c), one notes that d-excess is high near the surface, where ocean water is evaporated under non-equilibrium conditions (Aemisegger and Sjolte, 2018), and slightly decreases with altitude due to shallow cloud

610 processing (Thurnherr et al., 2021). Inside the lower part of the clouds, however, d-excess remains high, especially in the vicinity of strong updrafts (Fig. 14d). Additionally, as d-excess is higher below precipitating clouds (Fig. 14c at about -60.5° E and -58.2° E) than below non-precipitating clouds (Fig. 14c at about -55.9° E), partial evaporation of falling hydrometeors may additionally increase the d-excess in the subcloud layer vapour, which is then transported to cloud base in convective updrafts. In other words, the observed contrasts in d-excess between *cloudy* and *dry-warm* environments at cloud base, similar to q and

615 δ^2H, reflect mesoscale transport as well as the microphysical processes embedded in this flow.

The full distribution width of cloud base anomalies (black in Fig. 15) is explained by the contrast between the different cloud base environments. In the case of specific humidity and δ^2H, the *dry-warm* and *cloud-rain* categories are centred (median

value) outside the interquartile range of the full distribution, while it is the other way around for the d-excess. This applies to both datasets, the observations and the simulations. It means that the d-excess is not as strongly influenced by the mesoscale circulation associated with clouds as $\delta^2\text{H}$ and specific humidity. Other factors (e.g., spatial gradients as in Fig. 13d) play an equally or even larger role in the formation of d-excess anomalies.

We have a look at the mean vertical velocity (fluctuations; in case of the ATR data) inside the different cloud base environments to support the assumption that the *dry-warm* and *cloud/cloud-rain* categories represent different branches of the mesoscale overturning circulation (Fig. 15d). In the COSMO$_{\text{iso,1km}}$ data, the mean vertical velocity is negative in the *dry-warm* and *clear* environments and positive in the *cloud* and *cloud-rain* environments, as we would expect from an overturning circulation. Similar contrasts between the cloud base environments are found in the ATR data for the mean vertical velocity fluctuations. Note, that for the ATR data, we look at fluctuations relative to the flight-segment mean, because the accuracy of $0.1\,\text{m s}^{-1}$ of the aircraft's vertical motion measurements is not sufficient for estimating mean vertical winds (Lenschow et al., 2007).

5   Summary and conclusion

A combination of observational data from EUREC[4]A was used to evaluate COSMO$_{\text{iso}}$ simulations at different horizontal resolutions with a focus on processes at cloud base, which is the rooting level for cloud formation. This is the first isotope-enabled modelling study of shallow cumulus clouds in the trades with a storm-resolving model and in real large-scale meteorological conditions (i.e., applying a spectral nudging of the free-tropospheric horizontal winds). The detailed process-based evaluation focusing on shallow clouds was only possible thanks to the novel and unique observations collected during EUREC[4]A, in particular onboard the French ATR-42 aircraft.

The comparison to satellite images showed that COSMO$_{\text{iso}}$ is capable to distinguish between Fish clouds, Flower clouds, and more finely structured cloud patterns, such as Gravel and Sugar. Overall, if the model resolution is high, the simulated mesoscale cloud organisation agrees better with the satellite images than for simulations with lower resolution. The model differentiates between periods with high-intensity (20-24 January and 6-13 February 2020) and low-intensity rainfall (period in between), which are evident in satellite-based observations. For the major part of the simulated period, the model output matches the diel cycle of the satellite-based precipitation product and earlier studies (Vial et al., 2019). The flight-to-flight, as well as the in-flight variability of the isotope signals at cloud base, are well captured by the model (similar to earlier studies looking at the meso- to synoptic-scale variability in humidity and isotope variables; Dahinden et al., 2021; de Vries et al., 2022).

However, for some variables, COSMO$_{\text{iso}}$ has important biases. The liquid water content and cloud fraction at cloud base are overestimated but approach the observed values at higher spatial model resolution (for cloud fraction, this behaviour is expected and in agreement with earlier studies, i.e., Heim et al., 2021). Precipitation is spatially more concentrated in the simulations than in the satellite product. This local concentration of precipitation increases with higher model resolution. We find biases in humidity, temperature, and water vapour isotopes in all three simulations. COSMO$_{\text{iso}}$ produces a distinct cold-dry bias at low

levels that is consistently associated with too depleted vapour and too high values of the secondary isotope variable d-excess. A cold-dry bias was previously observed in other regional models such as HARMONIE (Van der Voort, 2021, humidity bias of similar magnitude; smaller bias in temperature) and AROME (Beucher et al., 2022, both biases of smaller magnitude), while too depleted vapour near the surface was found in COSMO$_{iso}$ simulations performed by Thurnherr et al. (2021, bias of approx.

655 2-4‰) over the Southern Ocean. Furthermore, COSMO$_{iso}$ shows too enriched vapour above the inversion (in agreement with Dahinden et al., 2021; de Vries et al., 2022) but no moist bias. This shows the additional value of isotope observations compared to specific humidity alone, and hints towards a misrepresentation of cloud top processes and their interaction with the dry free troposphere in the model. Different mechanisms that may cause individual biases were discussed and lead to the following recommendations for future research.

660 We recommend testing the role of turbulent mixing by adjusting the mixing length of the turbulence scheme. Mixing processes in the subcloud layer and at cloud base could be further investigated by combining turbulence and isotope observations from the ATR or by investigating the three-dimensional fields of total kinetic energy in the COSMO$_{iso,1km}$ data. The role of the fixed cloud droplet number, which is large compared to the observed spectrum, for the spatial distribution of cloud and rainwater should be assessed. Cloud top processes in the COSMO$_{iso}$ data could be investigated in more detail since near the

665 trade inversion isotopes might provide interesting additional information compared to specific humidity alone (formation of the stratiform outflow and sublimation of cloud droplets). Furthermore, we suggest repeating the simulations without applying spectral nudging to assess its potential role in the formation of the low-level cold-dry and depletion bias.

The motivation for the evaluation of the three COSMO$_{iso}$ simulations was to assess, whether they can be used to study the role of the atmospheric circulation at different scales for controlling the formation of shallow cumulus clouds in the

670 trades using isotopes. The realistic representation of cloud organisation patterns, cloud fraction, precipitation regimes and the temporal evolution of isotope variables at cloud base suggest that COSMO$_{iso,1km}$ can serve this purpose, while the simulations with coarser resolution are still useful for looking at the role of the larger-scale circulation as a driver for shallow cloud organisation. The identified model biases in humidity, temperature, and isotope signals in vapour are of minor concern because they can be eliminated by looking at anomalies, as shown in the feature-based evaluation (distinguishing between different

675 environments named *dry-warm*, *clear*, *cloud*, and *cloud-rain*) of COSMO$_{iso,1km}$. The feature-based evaluation revealed that the low-level biases in the COSMO$_{iso}$ simulations discussed above are not feature-specific but apply to all the different cloud base environments (cf. Supplement 3). In other words, the processes linked to the formation of these features are accurately reproduced by COSMO$_{iso}$.

The investigation of the isotope variability at cloud base identified mesoscale transport (overturning circulation associated

680 with clouds) and not local microphysical processes as the main control of δ^2H. The heaviest vapour with the highest d-excess was found inside clouds (directly linked to shallow convective upward transport and partial evaporation of falling hydrometeors), and the lightest vapour with the lowest d-excess inside dry-warm, clear-sky environments (directly linked to mesoscale subsidence), in the observations as well as in the COSMO$_{iso,1km}$ data. The clear relation between isotope signals and cloud base features reflecting the mesoscale circulation prompts the question whether isotope signals at cloud base contain

685   information on the strength of the cloud-relative overturning circulation. This question is addressed in the second part of this study by adopting a Lagrangian perspective.

Data availability. The ERA5 reanalysis dataset (Hersbach et al., 2020) can be downloaded from the ECMWF's official website (https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5, last access: 29 October 2022). The daily (Huffman et al., 2019b, https://doi.org/10.5067/GPM/IMERGDF/DAY/06, last access: 29 October 2022) and half-hourly (Huffman et al., 2019a, https://doi.org/10.5067/GPM/IMERG/3B-HH/06, last access: 29 October 2022) GPM IMERG data is available on NASA's Earthdata platform (https://disc.gsfc.nasa.gov/, last access: 6 April 2022). The six Suomi satellite images of the Earth Observing System Data and Information System (EOSDIS) are available in the Worldview Snapshots application (https://go.nasa.gov/3SOFwJw; https://go.nasa.gov/3fDXPCP; https://go.nasa.gov/3rsf1Oq; https://go.nasa.gov/3RvSxGA; https://go.nasa.gov/3e43ZvB; https://go.nasa.gov/3Cs4au5; last access: 3 October 2022). The aircraft-based observations, i.e., the CORE (CNRM/TRAMM et al., 2021), the BASTALIAS (Delanoë et al., 2021), the PMA (Coutris, 2021), the Picarro (Aemisegger et al., 2021a), and the JOANNE (George et al., 2021b) data can be obtained from the EUREC^{4}A data repository AERIS (https://eurec4a.aeris-data.fr, last access: 18 September 2022). The data from the GOES-16 visible channel (no. 2) can be downloaded from the EUREC^{4}A data catalogue (https://observations.ipsl.fr/thredds/catalog/EUREC4A/SATELLITES/GOES-E/catalog.html, last access: 18 September 2022). The ECHAM6-wiso are available upon request (contact: Martin Werner), and the COSMO$_{iso}$ simulations are published in the ETH research collection (Villiger and Aemisegger, 2022, note that the simulations have been extended and now provide data for the period from 20 January to 20 February 2020).

Author contributions. FA designed the project and acquired the funding. LV, with the support of FA, carried out the COSMO$_{iso}$ simulations, performed the data analysis, and wrote the paper, on which all the co-authors commented. LV, FA, and HW discussed the results and the structure of the paper in detail. SB initiated and coordinated the ATR observations. LV, MD, SB, ML, PEB, PC, PCo, JD, CF, AS, and FA carried out the observations onboard the ATR during EUREC^{4}A and assisted in their interpretation. FA, MD, SP were involved in the development and the technical setup of the COSMO$_{iso}$ simulations and assisted in their interpretation. LV, FA, ML, PEB discussed various possible definitions of warm-dry blobs at cloud base and their characteristic signature in the ATR observations. MW performed the ECHAM6-wiso simulation.

Competing interests. The authors have no competing interests to declare.

Acknowledgements. We thank Urs Beyerle (ETH Zürich) for the technical support regarding the cluster on which the model calculations were performed and Fabienne Dahinden for the help with the installation of the model on the cluster. The ATR aircraft was operated by SAFIRE, the French facility for airborne research, an infrastructure of the French National Center for Scientific Research (CNRS), Météo-France and the French National Center for Space Studies (CNES). We are grateful to SAFIRE for the realisation of the EUREC^{4}A airborne operations. We acknowledge the MeteoSwiss and the ECMWF for the access to the ERA5 reanalyses, NASA for the access to the satellite-based data,

and all those who were involved in the EUREC[4]A observations used in this publication. We are grateful to the two anonymous reviewers for

715 their constructive comments.

Financial support. LV received funding from the Swiss National Science Foundation (grant no. 188731). EUREC[4]A was funded by multiple national and international organisations, including the European Research Council (ERC; grant no. 694768), the Max Planck Society (MPG), the German Research Foundation (DFG), the German Meteorological Weather Service (DWD), the German Aerospace Center (DLR), and the National Centre for Space Studies CNES (EECLAT proposal). The cloud in situ data was collected using instruments from the French

720 Airborne measurement Platform, a facility partially funded by CNRS/INSU and CNES. SP acknowledges funding from the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft), project ID 441025101.

References

Aemisegger, F.: On the link between the North Atlantic storm track and precipitation deuterium excess in Reykjavik, Atmos. Sci. Lett., 19, 1–9, https://doi.org/10.1002/asl.865, 2018.

Aemisegger, F. and Sjolte, J.: A climatology of strong large-scale ocean evaporation events. Part II: Relevance for the deuterium excess signature of the evaporation flux, J. Climate, 31, 7313–7336, https://doi.org/10.1175/JCLI-D-17-0592.1, 2018.

Aemisegger, F., Spiegel, J. K., Pfahl, S., Sodemann, H., Eugster, W., and Wernli, H.: Isotope meteorology of cold front passages: A case study combining observations and modeling, Geophys. Res. Lett., 42, 5652–5660, https://doi.org/10.1002/2015GL063988, 2015.

Aemisegger, F., Dütsch, M., Rösch, M., and Villiger, L.: Calibrated stable water vapour isotope measurements on board the ATR-42 during EUREC4A [data set], AERIS, https://doi.org/10.25326/244, 2021a.

Aemisegger, F., Vogel, R., Graf, P., Dahinden, F., Villiger, L., Jansen, F., Bony, S., Stevens, B., and Wernli, H.: How Rossby wave breaking modulates the water cycle in the North Atlantic trade wind region, Weather Clim. Dynam., 2, 281–309, https://doi.org/10.5194/wcd-2-281-2021, 2021b.

Albright, A. L., Bony, S., Stevens, B., and Vogel, R.: Observed subcloud layer moisture and heat budgets in the trades, J. Atmos. Sci., 79, 2363–2385, https://doi.org/10.1175/JAS-D-21-0337.1, 2022.

Bailey, A., Nusbaumer, J., and Noone, D.: Precipitation efficiency derived from isotope ratios in water vapor distinguishes dynamical and microphysical influences on subtropical atmospheric constituents, J. Geophys. Res.-Atmos., 120, 9119–9137, https://doi.org/10.1002/2015JD023403, 2015.

Bailey, A., Aemisegger, F., Villiger, L., Los, S. A., Reverdin, G., Quiñones Meléndez, E., Acquistapace, C., Baranowski, D. B., Böck, T., Bony, S., Bordsdorff, T., Coffman, D., de Szoeke, S. P., Diekmann, C. J., Dütsch, M., Ertl, B., Galewsky, J., Henze, D., Przemyslaw, M., Noone, D., Quinn, P. K., Rösch, M., Schneider, A., Schneider, M., Speich, S., Bjorn, S., and Thompson, E. J.: Isotopic measurements in water vapor, precipitation, and seawater during EUREC4A, Earth Syst. Sci. Data, 15, 465–495, https://doi.org/10.5194/essd-15-465-2023, 2023.

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.

Beucher, F., Couvreux, F., Bouniol, D., Faure, G., Favot, F., Dauhut, T., and Ayet, A.: Process oriented evaluation of the oversea AROME configuration: focus on the representation of cloud organisation, Q. J. Roy. Meteor. Soc., pp. 2–37, https://doi.org/10.1002/qj.4354, 2022.

Blossey, P. N., Kuang, Z., and Romps, D. M.: Isotopic composition of water in the tropical tropopause layer in cloud-resolving simulations of an idealized tropical circulation, J. Geophys. Res.-Atmos., 115, 1–23, https://doi.org/10.1029/2010JD014554, 2010.

Bonne, J. L., Behrens, M., Meyer, H., Kipfstuhl, S., Rabe, B., Schönicke, L., Steen-Larsen, H. C., and Werner, M.: Resolving the controls of water vapour isotopes in the Atlantic sector, Nat. Commun., 10, https://doi.org/10.1038/s41467-019-09242-6, 2019.

Bony, S. and Dufresne, J. L.: Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models, Geophys. Res. Lett., 32, 1–4, https://doi.org/10.1029/2005GL023851, 2005.

Bony, S. and Stevens, B.: Measuring area-averaged vertical motions with dropsondes, J. Atmos. Sci., 76, 767–783, https://doi.org/10.1175/JAS-D-18-0141.1, 2019.

Bony, S., Risi, C., and Vimeux, F.: Influence of convective processes on the isotopic composition ($\delta18O$ and δD) of precipitation and water vapor in the tropics: 1. radiative-convective equilibrium and tropical ocean-global atmosphere-coupled ocean-atmosphere response experiment (TOGA-COARE), J. Geophys. Res.-Atmos., 113, 1–21, https://doi.org/10.1029/2008JD009942, 2008.

Bony, S., Stevens, B., Ament, F., Bigorre, S., Chazette, P., Crewell, S., Delanoë, J., Emanuel, K., Farrell, D., Flamant, C., Gross, S., Hirsch, L., Karstensen, J., Mayer, B., Nuijens, L., Ruppert, J. H., Sandu, I., Siebesma, P., Speich, S., Szczap, F., Totems, J., Vogel, R., Wendisch, M., and Wirth, M.: EUREC4A: A field campaign to elucidate the couplings between clouds, convection and circulation, Surv. Geophys., 38, 1529–1568, https://doi.org/10.1007/s10712-017-9428-0, 2017.

Bony, S., Schulz, H., Vial, J., and Stevens, B.: Sugar, gravel, fish, and flowers: Dependence of mesoscale patterns of trade-wind clouds on environmental conditions, Geophys. Res. Lett., 47, 1–9, https://doi.org/10.1029/2019GL085988, 2020.

Bony, S., Lothon, M., Delanoë, J., Coutris, P., Etienne, J.-C., Aemisegger, F., Albright, A. L., André, T., Bellec, H., Baron, A., Bourdinot, J.-F., Brilouet, P.-E., Bourdon, A., Canonici, J.-C., Caudoux, C., Chazette, P., Cluzeau, M., Cornet, C., Desbios, J.-P., Duchanoy, D., Flamant, C., Fildier, B., Gourbeyre, C., Guiraud, L., Jiang, T., Lainard, C., Le Gac, C., Lendroit, C., Lernould, J., Perrin, T., Pouvesle, F., Richard, P., Rochetin, N., Salaün, K., Schwarzenboeck, A., Seurat, G., Stevens, B., Totems, J., Touzé-Peiffer, L., Vergez, G., Vial, J., Villiger, L., and Vogel, R.: EUREC4A observations from the SAFIRE ATR42 aircraft, Earth Syst. Sci. Data, 14, 2021–2064, https://doi.org/10.5194/essd-14-2021-2022, 2022.

Cauquoin, A. and Werner, M.: High-resolution nudged isotope modeling with ECHAM6-Wiso: Impacts of Updated Model Physics and ERA5 Reanalysis Data, J. Adv. Model. Earth Sy., 13, 1–19, https://doi.org/10.1029/2021MS002532, 2021.

Cauquoin, A., Werner, M., and Lohmann, G.: Water isotopes - Climate relationships for the mid-Holocene and preindustrial period simulated with an isotope-enabled version of MPI-ESM, Clim. Past, 15, 1913–1937, https://doi.org/10.5194/cp-15-1913-2019, 2019.

Chazette, P., Totems, J., Baron, A., Flamant, C., and Bony, S.: Trade-wind clouds and aerosols characterized by airborne horizontal lidar measurements during the EUREC4A field campaign, Earth Syst. Sci. Data, 12, 2919–2936, https://doi.org/10.5194/essd-12-2919-2020, 2020.

Chazette, P., Flamant, C., Sodemann, H., Totems, J., Monod, A., Dieudonné, E., Baron, A., Seidl, A., Steen-Larsen, H. C., Doira, P., Durand, A., and Ravier, S.: Experimental investigation of the stable water isotope distribution in an Alpine lake environment (L-WAIVE), Atmos. Chem. Phys., 21, 10 911–10 937, https://doi.org/10.5194/acp-21-10911-2021, 2021.

Christner, E., Aemisegger, F., Pfahl, S., Werner, M., Cauquoin, A., Schneider, M., Hase, F., Barthlott, S., and Schädler, G.: The climatological impacts of continental surface evaporation, rainout, and subcloud processes on δD of water vapor and precipitation in Europe, J. Geophys. Res.-Atmos., 123, 4390–4409, https://doi.org/10.1002/2017JD027260, 2018.

CNRM/TRAMM, SAFIRE, and Laboratoire d'Aérologie: SAFIRE ATR42: Core Data 1Hz - V2 [data set], AERIS, https://doi.org/10.25326/298, 2021.

Coutris, P.: SAFIRE ATR42: PMA/cloud composite dataset [data set], AERIS, https://doi.org/10.25326/237, 2021.

Dahinden, F., Aemisegger, F., Wernli, H., Schneider, M., Diekmann, C. J., Ertl, B., Knippertz, P., Werner, M., and Pfahl, S.: Disentangling different moisture transport pathways over the eastern subtropical North Atlantic using multi-platform isotope observations and high-resolution numerical modelling, Atmos. Chem. Phys., 21, 16 319–16 347, https://doi.org/10.5194/acp-21-16319-2021, 2021.

Dauhut, T., Couvreux, F., Bouniol, D., Beucher, F., Volkmer, L., Pörtge, V., Schäfer, M., Ayet, A., Brilouet, P. E., Jacob, M., and Wirth, M.: Flower trade-wind clouds are shallow mesoscale convective systems, Q. J. Roy. Meteor. Soc., pp. 1–23, https://doi.org/10.1002/qj.4409, 2022.

de Vries, A. J., Aemisegger, F., Pfahl, S., and Wernli, H.: Stable water isotope signals in tropical ice clouds in the West African monsoon simulated with a regional convection-permitting model, Atmos. Chem. Phys., 22, 8863–8895, https://doi.org/10.5194/acp-22-8863-2022, 2022.

Delanoë, J., Protat, A., Vinson, J. P., Brett, W., Caudoux, C., Bertrand, F., du Chatelet, J. P., Hallali, R., Barthes, L., Haeffelin, M., and Dupont, J. C.: BASTA: A 95-GHz FMCW Doppler radar for cloud and fog studies, J. Atmos. Ocean. Tech., 33, 1023–1038, https://doi.org/10.1175/JTECH-D-15-0104.1, 2016.

Delanoë, J., Chazette, P., and Bony, S.: SAFIRE ATR42: BASTALIAS L2 dataset [data set] , AERIS, https://doi.org/10.25326/316, 2021.

800 Doms, G., Förstner, J., Heise, E., Herzog, H. J., Mironov, D., Raschendorfer, M., Reinhardt, T., Ritter, B., Schrodin, R., Schulz, J. P., and Vogel, G.: A description of the nonhydrostatic regional COSMO model. Part II: Physical parameterization, Tech. rep., Deutscher Wetterdienst, Offenbach, Germany, 2011.

Dütsch, M.: A new theoretical framework for parameterizing nonequilibrium fractionation during evaporation from the ocean [conference presentation], Workshop, 15-17 November, Water isotopes: from weather to climate, https://youtu.be/Cd28fK_1TUs, 2021.

805 Dütsch, M., Pfahl, S., and Sodemann, H.: The impact of nonequilibrium and equilibrium fractionation on two different deuterium excess definitions, J. Geophys. Res.-Atmos., 122, 12 732–12 746, https://doi.org/10.1002/2017JD027085, 2017.

Dyroff, C., Sanati, S., Christner, E., Zahn, A., Balzer, M., Bouquet, H., McManus, J. B., González-Ramos, Y., and Schneider, M.: Airborne in situ vertical profiling of HDO/H216O in the subtropical troposphere during the MUSICA remote sensing validation campaign, Atmos. Meas. Tech., 8, 2037–2049, https://doi.org/10.5194/amt-8-2037-2015, 2015.

810 Eirund, G. K., Drossaart van Dusseldorp, S., Brem, B. T., Dedekind, Z., Karrer, Y., Stoll, M., and Lohmann, U.: Aerosol–cloud–precipitation interactions during a Saharan dust event – A summertime case-study from the Alps, Q. J. Roy. Meteor. Soc., 148, 943–961, https://doi.org/10.1002/qj.4240, 2022.

Galewsky, J., Steen-Larsen, H. C., Field, R. D., Worden, J., Risi, C., and Schneider, M.: Stable isotopes in atmospheric water vapor and applications to the hydrologic cycle, Rev. Geophys., 54, 809–865, https://doi.org/10.1002/2015RG000512, 2016.

815 Gat, J. R.: Oxygen and hydrogen isotopes in the hydrologic cycle, Annu. Rev. Earth Pl. Sc., 24, 225–262, https://doi.org/10.1146/annurev.earth.24.1.225, 1996.

George, G., Stevens, B., Bony, S., Pincus, R., Fairall, C., Schulz, H., Kölling, T., Kalen, Q. T., Klingebiel, M., Konow, H., Lundry, A., Prange, M., and Radtke, J.: JOANNE: Joint dropsonde Observations of the Atmosphere in tropical North atlaNtic meso-scale Environments, Earth Syst. Sci. Data, 13, 5253–5272, https://doi.org/10.5194/essd-13-5253-2021, 2021a.

820 George, G., Stevens, B., Bony, S., Pincus, R., Fairall, C., Schulz, H., Kölling, T., Kalen, T. Q. T., Klingebiel, M., Konow, H., Lundry, A., Prange, M., and Radtke, J.: JOANNE: Joint dropsonde Observations of the Atmosphere in tropical North atlaNtic meso-scale Environments [data set], AERIS, https://doi.org/10.25326/221, 2021b.

George, G., Stevens, B., Bony, S., Vogel, R., and Naumann, A. K.: Widespread shallow mesoscale circulations observed in the trades, Nat. Geosci., 16, 584–589, https://doi.org/https://doi.org/10.1038/s41561-023-01215-1, 2023.

825 Graf, P., Wernli, H., Pfahl, S., and Sodemann, H.: A new interpretative framework for below-cloud effects on stable water isotopes in vapour and rain, Atmos. Chem. Phys., 19, 747–765, https://doi.org/10.5194/acp-19-747-2019, 2019.

Guilpart, E., Vimeux, F., Evan, S., Brioude, J., Metzger, J. M., Barthe, C., Risi, C., and Cattani, O.: The isotopic composition of near-surface water vapor at the Maïdo observatory (Reunion Island, southwestern Indian Ocean) documents the controls of the humidity of the subtropical troposphere, J. Geophys. Res.-Atmos., 122, 9628–9650, https://doi.org/10.1002/2017JD026791, 2017.

830 Hanley, K. E., Plant, R. S., Stein, T. H. M., Hogan, R. J., Nicol, J. C., Lean, H. W., Halliwell, C., and Clark, P. A.: Mixing-length controls on high-resolution simulations of convective storms, Q. J. Roy. Meteor. Soc., 141, 272–284, https://doi.org/10.1002/qj.2356, 2015.

Heim, C., Hentgen, L., Ban, N., and Schär, C.: Inter-model variability in convection-resolving simulations of subtropical marine low clouds, J. Meteorol. Soc. Jpn., 99, 1271–1295, https://doi.org/10.2151/jmsj.2021-062, 2021.

Hentgen, L.: Clouds in convection-resolving climate simulations over Europe and the tropical Atlantic, Ph.D. thesis, ETH Zürich, https://doi.org/10.3929/ethz-b-000403020, 2019.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J. N.: The ERA5 global reanalysis, Q. J. Roy. Meteor. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Holton, J. R. and Hakim, G. J.: Tropical dynamics, in: Introduction to dynamic meterology, edited by Holton, J. R. and Hakim, G. J., chap. 11, pp. 377–411, Academic Press, 5th edn., https://doi.org/10.1016/B978-0-12-384866-6.00011-8, 2013.

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 x 0.1 degree V06 [data set], Goddard Earth Sciences Data and Information Services Center (GES DISC), https://doi.org/10.5067/GPM/IMERG/3B-HH/06, 2019a.

Huffman, G. J., Stocker, E. F., Bolvin, D. T., Nelkin, E. J., and Tan, J.: GPM IMERG final precipitation L3 1 day 0.1 degree x 0.1 degree V06 [data set], Goddard Earth Sciences Data and Information Services Center (GES DISC), https://doi.org/10.5067/GPM/IMERGDF/DAY/06, 2019b.

International Atomic Energy Agency: Reference Sheet for VSMOW2 and SLAP2 international measurement standards., IAEA, Vienna, 2017.

Jeevanjee, N.: Vertical velocity in the gray zone, J. Adv. Model. Earth Sy., 9, 2304–2316, https://doi.org/10.1002/2017MS001059, 2017.

Konow, H., Ewald, F., George, G., Jacob, M., Klingebiel, M., Kölling, T., Luebke, A. E., Mieslinger, T., Pörtge, V., Radtke, J., Schäfer, M., Schulz, H., Vogel, R., Wirth, M., Bony, S., Crewell, S., Ehrlich, A., Forster, L., Giez, A., Gödde, F., Groß, S., Gutleben, M., Hagen, M., Hirsch, L., Jansen, F., Lang, T., Mayer, B., Mech, M., Prange, M., Schnitt, S., Vial, J., Walbröl, A., Wendisch, M., Wolf, K., Zinner, T., Zöger, M., Ament, F., and Stevens, B.: EUREC4A's HALO, Earth Syst. Sci. Data, 13, 5545–5563, https://doi.org/10.5194/essd-13-5545-2021, 2021.

Lance, S., Brock, C. A., Rogers, D., and Gordon, J. A.: Water droplet calibration of the Cloud Droplet Probe (CDP) and in-flight performance in liquid, ice and mixed-phase clouds during ARCPAC, Atmos. Meas. Tech., 3, 1683–1706, https://doi.org/10.5194/amt-3-1683-2010, 2010.

Lawson, R. P., O'Connor, D., Zmarzly, P., Weaver, K., Baker, B., Mo, Q., and Jonsson, H.: The 2D-S (stereo) probe: Design and preliminary tests of a new airborne, high-speed, high-resolution particle imaging probe, J. Atmos. Ocean. Tech., 23, 1462–1477, https://doi.org/10.1175/JTECH1927.1, 2006.

Lenschow, D. H., Savic-Jovcic, V., and Stevens, B.: Divergence and vorticity from aircraft air motion measurements, J. Atmos. Ocean. Tech., 24, 2062–2072, https://doi.org/10.1175/2007JTECHA940.1, 2007.

Merlivat, L. and Jouzel, J.: Global climatic interpretation of the deuterium-oxygen 18 relationship for precipitation, J. Geophys. Res.-Oceans, 84, 5029–5030, https://doi.org/10.1029/JC084iC08p05029, 1979.

Naumann, A. K., Stevens, B., Hohenegger, C., and Mellado, J. P.: A conceptual model of a shallow circulation induced by prescribed low-level radiative cooling, J. Atmos. Sci., 74, 3129–3144, https://doi.org/10.1175/JAS-D-17-0030.1, 2017.

NOAA: Visible Infrared Imaging Radiometer Suite (VIIRS) Sensor Data Record (SDR) User's Guide. Version 1.3., Tech. rep., U.S. Department of Commerce, 2017.

Noone, D.: Pairing measurements of the water vapor isotope ratio with humidity to deduce atmospheric moistening and dehydration in the tropical midtroposphere, J. Climate, 25, 4476–4494, https://doi.org/10.1175/JCLI-D-11-00582.1, 2012.

Noone, D., Galewsky, J., Sharp, Z. D., Worden, J., Barnes, J., Baer, D., Bailey, A., Brown, D. P., Christensen, L., Crosson, E., Dong, F., Hurley, J. V., Johnson, L. R., Strong, M., Toohey, D., Van Pelt, A., and Wright, J. S.: Properties of air mass mixing and humidity in the subtropics from measurements of the D/H isotope ratio of water vapor at the Mauna Loa Observatory, J. Geophys. Res.-Atmos., 116, 1–18, https://doi.org/10.1029/2011JD015773, 2011.

Pfahl, S. and Wernli, H.: Air parcel trajectory analysis of stable isotopes in water vapor in the eastern Mediterranean, J. Geophys. Res.-Atmos., 113, 1–16, https://doi.org/10.1029/2008JD009839, 2008.

Pfahl, S. and Wernli, H.: Lagrangian simulations of stable isotopes in water vapor: An evaluation of nonequilibrium fractionation in the Craig-Gordon model, J. Geophys. Res.-Atmos., 114, 1–12, https://doi.org/10.1029/2009JD012054, 2009.

Pfahl, S., Wernli, H., and Yoshimura, K.: The isotopic composition of precipitation from a winter storm-a case study with the limited-area model COSMOiso, Atmos. Chem. Phys., 12, 1629–1648, https://doi.org/10.5194/acp-12-1629-2012, 2012.

Risi, C., Bony, S., and Vimeux, F.: Influence of convective processes on the isotopic composition (δ18O and δD) of precipitation and water vapor in the tropics: 2. Physical interpretation of the amount effect, J. Geophys. Res.-Atmos., 113, 1–12, https://doi.org/10.1029/2008JD009943, 2008.

Risi, C., Galewsky, J., Reverdin, G., and Brient, F.: Controls on the water vapor isotopic composition near the surface of tropical oceans and role of boundary layer mixing processes, Atmos. Chem. Phys., 19, 12 235–12 260, https://doi.org/10.5194/acp-19-12235-2019, 2019.

Risi, C., Muller, C., and Blossey, P.: What controls the water vapor isotopic composition near the surface of tropical oceans? Results from an analytical model constrained by large-eddy simulations, J. Adv. Model. Earth Sy., 12, 1–27, https://doi.org/10.1029/2020MS002106, 2020.

Risi, C., Muller, C., and Blossey, P.: Rain evaporation, snow melt, and entrainment at the heart of water vapor isotopic variations in the tropical troposphere, according to large-eddy simulations and a two-column model, J. Adv. Model. Earth Sy., 13, 1–28, https://doi.org/10.1029/2020MS002381, 2021.

Salathé, E. P. and Hartmann, D. L.: A trajectory analysis of tropical upper-tropospheric moisture and convection, J. Climate, 10, 2533–2547, https://doi.org/10.1175/1520-0442(1997)010<2533:ATAOTU>2.0.CO;2, 1997.

Salmon, O. E., Welp, L. R., Baldwin, M. E., Hajny, K. D., Stirm, B. H., and Shepson, P. B.: Vertical profile observations of water vapor deuterium excess in the lower troposphere, Atmos. Chem. Phys., 19, 11 525–11 543, https://doi.org/10.5194/acp-19-11525-2019, 2019.

Savazzi, A. C. M., Nuijens, L., Sandu, I., George, G., and Bechtold, P.: The representation of winds in the lower troposphere in ECMWF forecasts and reanalyses during the EUREC4A field campaign, Atmos. Chem. Phys., 22, 13 049–13 066, https://doi.org/10.5194/acp-22-13049-2022, 2022.

Schlemmer, L. and Hohenegger, C.: The formation of wider and deeper clouds as a result of cold-pool dynamics, J. Atmos. Sci., 71, 2842–2858, https://doi.org/10.1175/JAS-D-13-0170.1, 2014.

Schneider, T., Teixeira, J., Bretherton, C. S., Brient, F., Pressel, K. G., Schär, C., and Siebesma, A. P.: Climate goals and computing the future of clouds, Nat. Clim. Change, 7, 3–5, https://doi.org/10.1038/nclimate3190, 2017.

Schubert-Frisius, M., Feser, F., von Storch, H., and Rast, S.: Optimal spectral nudging for global dynamic downscaling, Mon. Weather Rev., 145, 909–927, https://doi.org/10.1175/MWR-D-16-0036.1, 2017.

Schulz, H.: C3ONTEXT: A Common Consensus on Convective OrgaNizaTion during the EUREC4A eXperimenT, Earth Syst. Sci. Data, 14, 1233–1256, https://doi.org/10.5194/essd-14-1233-2022, 2022.

Schulz, H. and Stevens, B.: On the representation of shallow convection in the trades by large-domain, hecto-meter, large-eddy simulations, J. Adv. Model. Earth Sy. [preprint], https://eartharxiv.org/repository/view/4986/, 2023.

Sodemann, H., Aemisegger, F., Pfahl, S., Bitter, M., Corsmeier, U., Feuerle, T., Graf, P., Hankers, R., Hsiao, G., Schulz, H., Wieser, A., and Wernli, H.: The stable isotopic composition of water vapour above Corsica during the HyMeX SOP1 campaign: Insight into vertical mixing processes from lower-tropospheric survey flights, Atmos. Chem. Phys., 17, 6125–6151, https://doi.org/10.5194/acp-17-6125-2017, 2017.

Steen-Larsen, H. C., Sveinbjörnsdottir, A. E., Jonsson, T., Ritter, F., Bonne, J., Masson-Delmotte, V., Sodemann, H., Blunier, T., Dahl-Jensen, D., and Vinther, B. M.: Moisture sources and synoptic to seasonal variability of North Atlantic water vapor isotopic composition, J. Geophys. Res.-Atmos., 120, 5757–5774, https://doi.org/10.1002/2015JD023234, 2015.

Steppeler, J., Doms, G., Schättler, U., Bitzer, H. W., Gassmann, A., Damrath, U., and Gregoric, G.: Meso-gamma scale forecasts using the nonhydrostatic model LM, Meteorol. Atmos. Phys., 82, 75–96, https://doi.org/10.1007/s00703-001-0592-9, 2003.

Stevens, B., Bony, S., Brogniez, H., Hentgen, L., Hohenegger, C., Kiemle, C., L'Ecuyer, T. S., Naumann, A. K., Schulz, H., Siebesma, P. A., Vial, J., Winker, D. M., and Zuidema, P.: Sugar, gravel, fish and flowers: Mesoscale cloud patterns in the trade winds, Q. J. Roy. Meteor. Soc., 146, 141–152, https://doi.org/10.1002/qj.3662, 2020.

Stevens, B., Bony, S., Farrell, D., Ament, F., Blyth, A., Fairall, C., Karstensen, J., Quinn, P., Speich, S., Aemisegger, F., Albright, A., Bodenschatz, E., Chewitt-Lucas, R., Delanoë, J., Ewald, F., Forde, M., George, G., Hausold, A., Hagen, M., Hirsch, L., Jansen, F., Kinne, S., Klocke, D., Kölling, T., Konow, H., Mohr, W., Naumann, A., Nuijens, L., Pincus, R., Reverdin, G., Roberts, G., Schnitt, S., Sullivan, P., Touzé-Peiffer, L., Vial, J., Vogel, R., Acquistapace, C., Alexander, N., Alves, L., Arixi, S., Asmath, H., Bagheri, G., Bailey, A., Baranowski, D., Baron, A., Barrau, S., Barrett, P., Behrendt, A., Bellenger, H., Bendinger, A., Beucher, F., Bigorre, S., Blossey, P., Bock, O., Bosser, P., Bourras, D., Bouruet-Aubertot, P., Bower, K., Branger, H., Brennek, M., Brewer, A., Brügmann, B., Buehler, S., Burke, E., Burton, R., Böing, S., Calmer, R., Carton, X., Cato, G., Charles, J., Chazette, P., Chen, Y., Choularton, T., Chuang, P., Clarke, S., Coe, H., Cornet, C., Coutris, P., Couvreux, F., Crewell, S., Cronin, T., Cui, Z., Cuypers, Y., Daley, A., Damerell, G., Dauhut, T., Denby, L., Deneke, H., Douet, V., Drushka, K., Dütsch, M., Ehrlich, A., Emanuel, K., Emmanouilidis, A., Etienne-Leblanc, S., Faure, G., Feingold, G., Ferrero, L., Fildier, B., Fix, A., Flamant, C., Flatau, P., Foltz, G., Gadian, A., Galewsky, J., Gallagher, M., Gallimore, P., Gaston, C., Gentemann, C., Geyskens, N., Giez, A., Gourbeyre, C., Gross, S., Grosz, R., Gutleben, M., Güttler, J., Hall, K., Harris, G., Helfer, K., Henze, D., Herbert, C., Heywood, K., Holanda, B., Ibanez-Landeta, A., Intrieri, J., Iyer, S., Jacob, M., Julien, F., Kalesse, H., Kazil, J., Kellman, A., Kirchner, U., Klingebiel, M., Kremper, L., Kretzschmar, J., Krüger, O., Kurz, A., Körner, M., Lachlan-Cope, T., Laing, A., Landschützer, P., Lang, T., Lange, D., Lange, I., Laplace, C., Laxenaire, R., Leandro, M., Lefevre, N., Lenschow, D., Li, Q., Lloyd, G., Los, S., Losi, N., Lothon, M., Lovell, O., Luneau, C., Makuch, P., Malinowski, S., Manta, G., Marinou, E., Marsden, N., Matthieu, L., Maury, N., Mayer, B., Mayers-Als, M., Mazel, C., Mcgeary, W., Mcwilliams, J., Mech, M., Mehlmann, M., Meroni, A., Mieslinger, T., Minikin, A., Muller, C., Musat, I., Möller, G., Napoli, A., Neuberger, A., Noone, D., Nordsiek, F., Nowak, J., Olivier, L., Oswald, L., Parker, D., Peck, C., Person, R., Plueddemann, A., Pologne, L., Posyniak, M., Prange, M., Pöhlker, M., Pöhlker, C., Pörtge, V., Pöschl, U., Radtke, J., Ramage, K., Reimann, J., Renault, L., Reus, K., Reyes, A., Ribbe, J., Ringel, M., Ritschel, M., Rocha, C., Rochetin, N., Rollo, C., Röttenbacher, J., Saffin, L., Sandiford, S., Sandu, I., Schemann, V., Schirmacher, I., Schlenczek, O., Schmidt, J., Schröder, M., Schulz, H., Schwarzenboeck, A., Schäfer, M., Sealy, A., Serikov, I., Shohan, S., Siddle, E., Siebesma, A., Späth, F., Stephan, C., Stolla, M., Szkółka, W., Tarot, S., Tetoni, E., Thompson, E., Thomson, J., Tomassini, L., Totems, J., Villiger, L., Walther, A., Webber, B., Wendisch, M., Whitehall, S., Wiltshire, A., Wing, A., Wirth, M., Wiskandt, J., Wolf, K., Worbes, L., Wright, E., Wulfmeyer, V.,

Young, S., Zhang, D., Zhang, C., Ziemen, F., Zinner, T., Zuidema, P., and Zöger, M.: EUREC4A, Earth Syst. Sci. Data, 13, 4067–4119, https://doi.org/10.5194/essd-13-4067-2021, 2021.

Sun, J., Zhang, K., Wan, H., Ma, P., Tang, Q., and Zhang, S.: Impact of nudging strategy on the climate representativeness and hindcast skill of constrained EAMv1 simulations, J. Adv. Model. Earth Sy., 11, 3911–3933, https://doi.org/10.1029/2019MS001831, 2019.

950 Theunert, F. and Seifert, A.: Simulation studies of shallow convection with the convection-resolving version of DWD Lokal-Modell, COSMO Newsl., 6, 121–128, 2006.

Thurnherr, I., Kozachek, A., Graf, P., Weng, Y., Bolshiyanov, D., Landwehr, S., Pfahl, S., Schmale, J., Sodemann, H., Steen-Larsen, H. C., Toffoli, A., Wernli, H., and Aemisegger, F.: Meridional and vertical variations of the water vapour isotopic composition in the marine boundary layer over the Atlantic and Southern Ocean, Atmos. Chem. Phys., 20, 5811–5835, https://doi.org/10.5194/acp-20-5811-2020, 955 2020.

Thurnherr, I., Hartmuth, K., Jansing, L., Gehring, J., Boettcher, M., Gorodetskaya, I., Werner, M., Wernli, H., and Aemisegger, F.: The role of air–sea fluxes for the water vapour isotope signals in the cold and warm sectors of extratropical cyclones over the Southern Ocean, Weather Clim. Dynam., 2, 331–357, https://doi.org/10.5194/wcd-2-331-2021, 2021.

Tiedtke, M.: A comprehensive mass flux scheme for cumulus parameterization in large-scale models, Mon. Weather Rev., 117, 1779–1800, 960 https://doi.org/10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2, 1989.

Tompkins, A. M.: Organization of tropical convection in low vertical wind shears: The role of cold pools, J. Atmos. Sci., 58, 1650–1672, https://doi.org/10.1175/1520-0469(2001)058<1650:OOTCIL>2.0.CO;2, 2001.

Torri, G.: On the isotopic composition of cold pools in radiative-convective equilibrium, J. Geophys. Res.-Atmos., 126, 1–20, https://doi.org/10.1029/2020JD033139, 2021.

965 Torri, G. and Kuang, Z.: Rain evaporation and moist patches in tropical boundary layers, Geophys. Res. Lett., 43, 9895–9902, https://doi.org/10.1002/2016GL070893, 2016a.

Torri, G. and Kuang, Z.: A Lagrangian study of precipitation-driven downdrafts, J. Atmos. Sci., 73, 839–854, https://doi.org/10.1175/JAS-D-15-0222.1, 2016b.

Torri, G., Kuang, Z., and Tian, Y.: Mechanisms for convection triggering by cold pools, Geophys. Res. Lett., 42, 1943–1950, 970 https://doi.org/10.1002/2015GL063227, 2015.

Torri, G., Ma, D., and Kuang, Z.: Stable water isotopes and large-scale vertical motions in the tropics, J. Geophys. Res.-Atmos., 122, 3703–3717, https://doi.org/10.1002/2016JD026154, 2017.

Touzé-Peiffer, L., Vogel, R., and Rochetin, N.: Cold pools observed during EUREC4A: detection and characterization of atmospheric soundings, J. Appl. Meteorol. Clim., 61, 593–610, https://doi.org/10.1175/JAMC-D-21-0048.1, 2022.

975 Van der Voort, F.: Evaluation of the representation of subtropical marine clouds in the numerical weather prediction model HARMONIE., 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.

Vial, J., Bony, S., Dufresne, J. L., and Roehrig, R.: Coupling between lower-tropospheric convective mixing and low-level clouds: Physical 980 mechanisms and dependence on convection scheme, J. Adv. Model. Earth Sy., 8, 1892–1911, https://doi.org/10.1002/2016MS000740, 2016.

Vial, J., Vogel, R., Bony, S., Stevens, B., Winker, D. M., Cai, X., Hohenegger, C., Naumann, A. K., and Brogniez, H.: A new look at the daily cycle of tradewind cumuli, J. Adv. Model. Earth Sy., 11, 3148–3166, https://doi.org/10.1029/2019ms001746, 2019.

Vial, J., Vogel, R., and Schulz, H.: On the daily cycle of mesoscale cloud organization in the winter trades, Q. J. Roy. Meteor. Soc., 147, 2850–2873, https://doi.org/10.1002/qj.4103, 2021.

Villiger, L.: Large-scale circulation drivers and stable water isotope characteristics of shallow clouds over the tropical North Atlantic, Ph.D. thesis, ETH Zürich, https://doi.org/10.3929/ethz-b-000586270, 2022.

Villiger, L. and Aemisegger, F.: Numerical weather simulation using COSMOiso over the tropical North Atlantic in January and February 2020 in the context of EUREC4A, ETH Research Collection, https://doi.org/10.3929/ethz-b-000584213, 2022.

Villiger, L., Wernli, H., Boettcher, M., Hagen, M., and Aemisegger, F.: Lagrangian formation pathways of moist anomalies in the trade-wind region during the dry season: two case studies from EUREC4A, Weather Clim. Dynam., 3, 59–88, https://doi.org/10.5194/wcd-3-59-2022, 2022.

Vogel, R., Konow, H., Schulz, H., and Zuidema, P.: A climatology of trade-wind cumulus cold pools and their link to mesoscale cloud organization, Atmos. Chem. Phys., 21, 16 609–16 630, https://doi.org/10.5194/acp-21-16609-2021, 2021.

Von Storch, H., Langenberg, H., and Feser, F.: A spectral nudging technique for dynamical downscaling purposes, Mon. Weather Rev., 128, 3664–3673, https://doi.org/10.1175/1520-0493(2000)128<3664:ASNTFD>2.0.CO;2, 2000.

Wehrli, K., Guillod, B. P., Hauser, M., Leclair, M., and Seneviratne, S. I.: Assessing the dynamic versus thermodynamic origin of climate model biases, Geophys. Res. Lett., 45, 8471–8479, https://doi.org/10.1029/2018GL079220, 2018.

Wood, R.: Stratocumulus clouds, Mon. Weather Rev., 140, 2373–2423, https://doi.org/10.1175/MWR-D-11-00121.1, 2012.

Wyngaard, J. C.: Toward numerical modeling in the "Terra Incognita", J. Atmos. Sci., 61, 1816–1826, https://doi.org/10.1175/1520-0469(2004)061<1816:TNMITT>2.0.CO;2, 2004.

Zannoni, D., Steen-Larsen, H. C., Peters, A. J., Wahl, S., Sodemann, H., and Sveinbjörnsdóttir, A. E.: Non-Equilibrium Fractionation Factors for D/H and 18O/16O During Oceanic Evaporation in the North-West Atlantic Region, J. Geophys. Res.-Atmos., 127, 1–18, https://doi.org/10.1029/2022JD037076, 2022.

Zelinka, M. D., Randall, D. A., Webb, M. J., and Klein, S. A.: Clearing clouds of uncertainty, Nat. Clim. Change, 7, 674–678, https://doi.org/10.1038/nclimate3402, 2017.

Appendix A: Description of supplement material

This paper is supplemented by the following four documents:

Supplement 1 contains an explanatory sketch of the cloud fraction's resolution dependency and further illustrates the spatial distribution of clouds as in Fig. 4, but for every day from 20 January to 13 February 2020. With this, we want to create transparency in how well the mesoscale cloud organisation patterns are reproduced by the simulations beyond the six time steps shown in the paper. Additionally, we include a version of the figures that displays a larger geographical domain than shown in Fig. 4 with the goal of informing about the cloud organisation on the large scale.

Supplement 2 provides further evaluation analyses addressing the horizontal wind components and the mesoscale vertical velocities at cloud base. The analysis should give an idea of how much the dynamical variables are constrained by the spectral nudging.

Supplement 3 informs about the number of identified *dry-warm* cloud base grid points in COSMO$_{iso,1km}$. Additionally, it contains further characterisations of the different cloud base environments (alternative version of Fig. 15). First, the environ-

ments' anomaly distributions are shown in their full widths (instead of the 25-75-percentile and 10-90-percentile range as in Fig. 15) to illustrate their mutual overlap. Second, the absolute values (instead of anomalous values as in Fig. 15) are shown to demonstrate that all cloud base environments in the simulations are characterised by a bias compared to the ATR observations.

Supplement 4 compares $COSMO_{iso}$ simulations with different formulations of the non-equilibrium fractionation factors. Namely, the $COSMO_{iso,5km}$ and $COSMO_{iso,1km}$ simulations described in the paper (using the smooth regime of Merlivat and Jouzel, 1979) and two $COSMO_{iso}$ simulations, which use the recently developed formulation by Dütsch (2021). The analysis shows that the choice of the non-equilibrium fraction factors clearly influences the simulated d-excess in vapour.
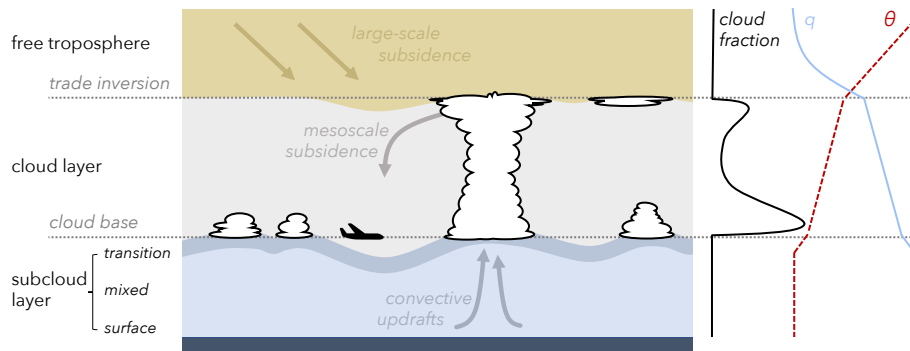
Figure 1. Schematic of the different atmospheric layers in the troposphere over the tropical Atlantic with the ATR aircraft flying at cloud base together with idealised profiles of cloud fraction (black), specific humidity (blue) and potential temperature (red dashed). Based on the ATR measurement time series shown in Fig. 12, we assume that the cloud base/subcloud layer top is characterised by an uneven topography instead of being a flat surface. See the text for the description of the individual layers.
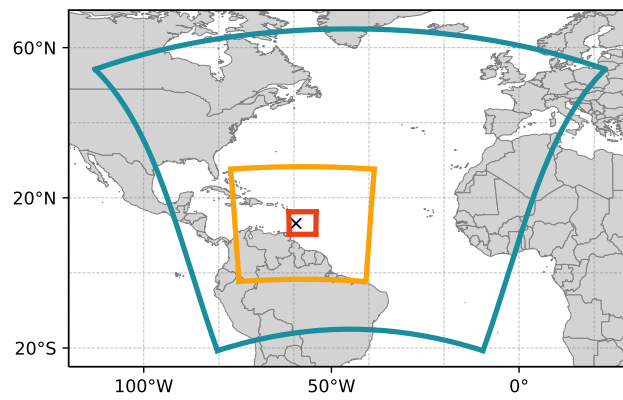
Figure 2. Domain boundaries of the three COSMO$_{iso}$ simulations, COSMO$_{iso,10km}$ (teal), COSMO$_{iso,5km}$ (yellow), and COSMO$_{iso,1km}$ (red). The location of Barbados is indicated by a black cross.
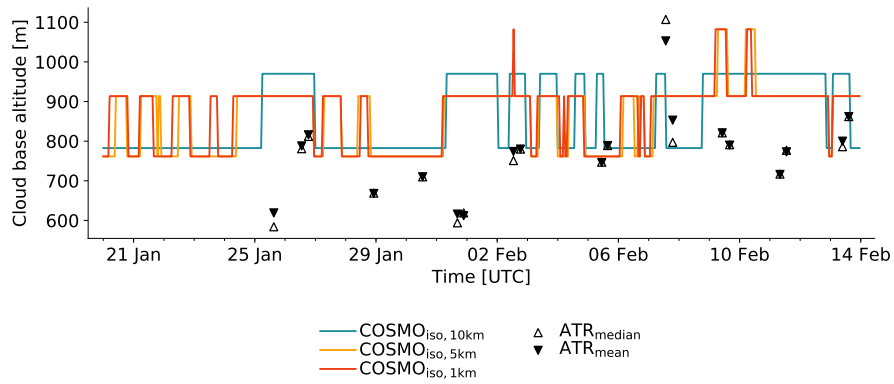
Figure 3. Time series of cloud base COSMO$_{iso}$ model level and ATR flight altitude. Shown are the median values of the grid points in the domain 54.5-61° W, 11-16° N at every hourly time step for COSMO$_{iso,10km}$ (teal), COSMO$_{iso,5km}$ (yellow), and COSMO$_{iso,1km}$ (red), as well as the median (empty black markers) and mean (filled black markers) value of each ATR flight.
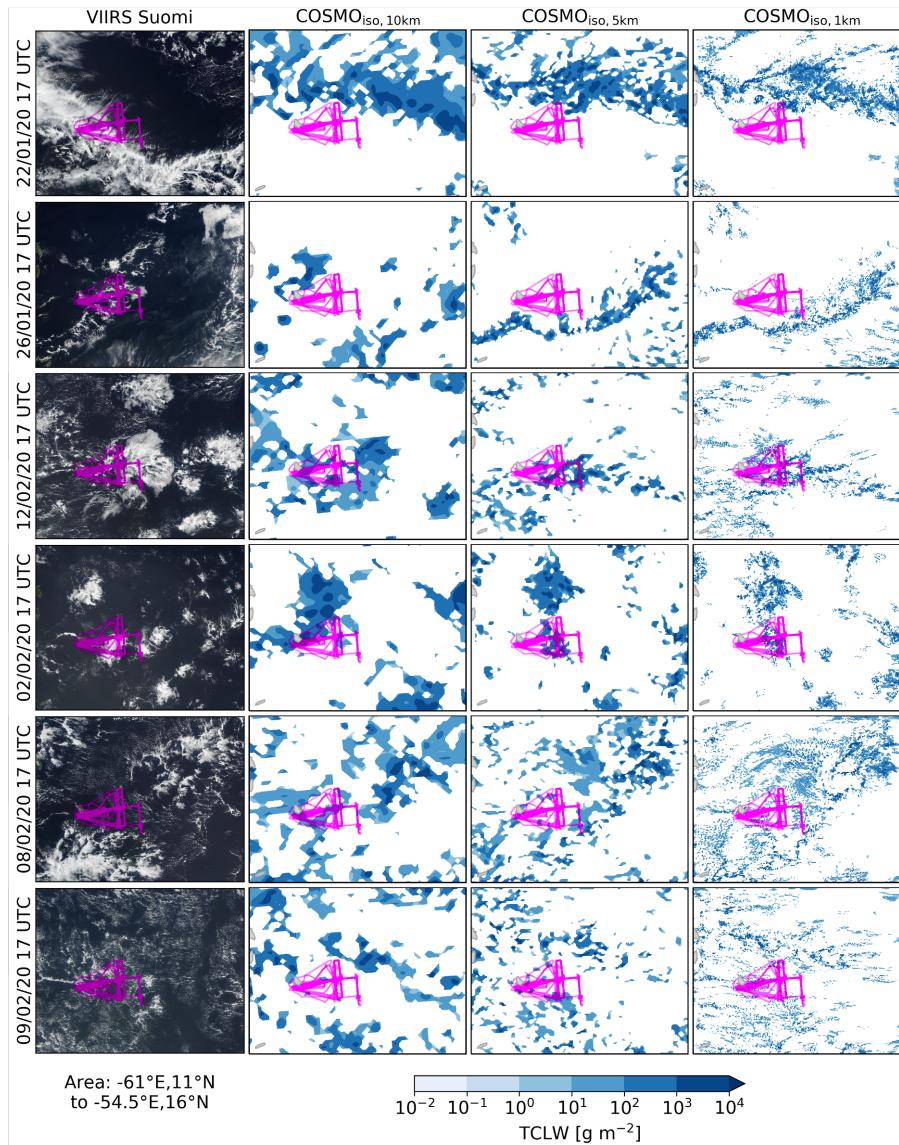
Figure 4. Spatial organisation of the clouds in the domain 54.5-61° W, 11-16° N from four datasets (columns) and on six dates (rows). The first column shows the image from the VIIRS instrument onboard the polar-orbiting Suomi-NPP satellite (approx. equatorial crossing time 17:30 UTC). The remaining columns show the total column cloud liquid water (TCLW) in the three COSMO$_{iso}$ simulations at 17:00 UTC. The dates are chosen according to the displayed cloud species, with large/small Fish clouds on 22/26 January 2020, large/small Flower clouds on 12/2 February 2020, and Gravel-Sugar/Sugar clouds on 8/9 February 2020. The track from all ATR flights is overlayed in pink. Other examples are shown in Supplement 1.
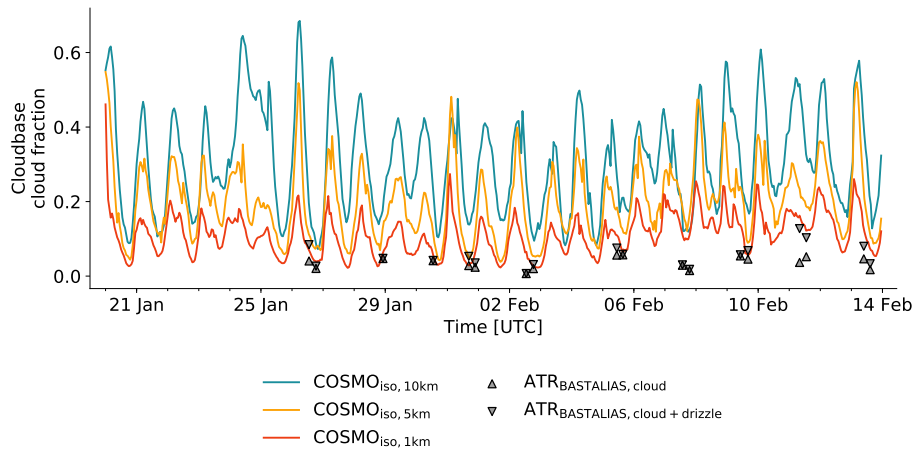
Figure 5. Time series of cloud fraction at cloud base for COSMO$_{iso,10km}$ (teal), COSMO$_{iso,5km}$ (yellow), COSMO$_{iso,1km}$ (red) and the ATR (black markers). For COSMO$_{iso}$, the cloud fraction is calculated as the fraction of cloud base grid points in the domain 54.5-61° W, 11-16° N with cloud water content exceeding $10\,\mathrm{mg\,kg^{-1}}$. The number of grid points in the considered domain are 3221 for COSMO$_{iso,10km}$, 12632 for COSMO$_{iso,5km}$, and 316028 for COSMO$_{iso,1km}$.
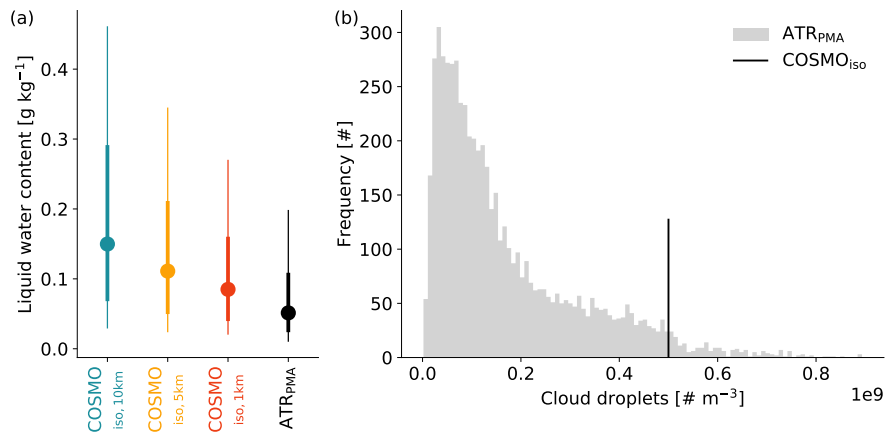
Figure 6. Distributions at cloud base of (a) liquid water content (LWC) from COSMO$_{iso,10km}$ (teal), COSMO$_{iso,5km}$ (yellow), COSMO$_{iso,1km}$ (red), and ATR$_{PMA}$ (black); and (b) cloud droplet numbers from COSMO$_{iso}$ (black) and ATR$_{PMA}$ (grey). For COSMO$_{iso}$, the LWC is calculated as the sum of cloud water content (CWC) and rainwater content (RWC; i.e., LWC = CWC + RWC). Shown are (a) the 10-90-percentile range (thin vertical line), 25-75-percentile range (thick vertical line), and the median (marker) of all cloud base data points with LWC> $10\,\text{mg}\,\text{kg}^{-1}$. For the ATR$_{PMA}$ dataset, all flights (RF02-RF20) are taken into account, and for the COSMO$_{iso}$, the data in the domain 54.5-61° W, 11-16° N from the hourly time steps that are closest to the ATR flights. In (b), the droplet number distribution from the ATR$_{PMA}$ dataset during time steps flagged as cloud is shown together with the constant cloud droplet number of COSMO$_{iso}$'s one-moment cloud scheme.
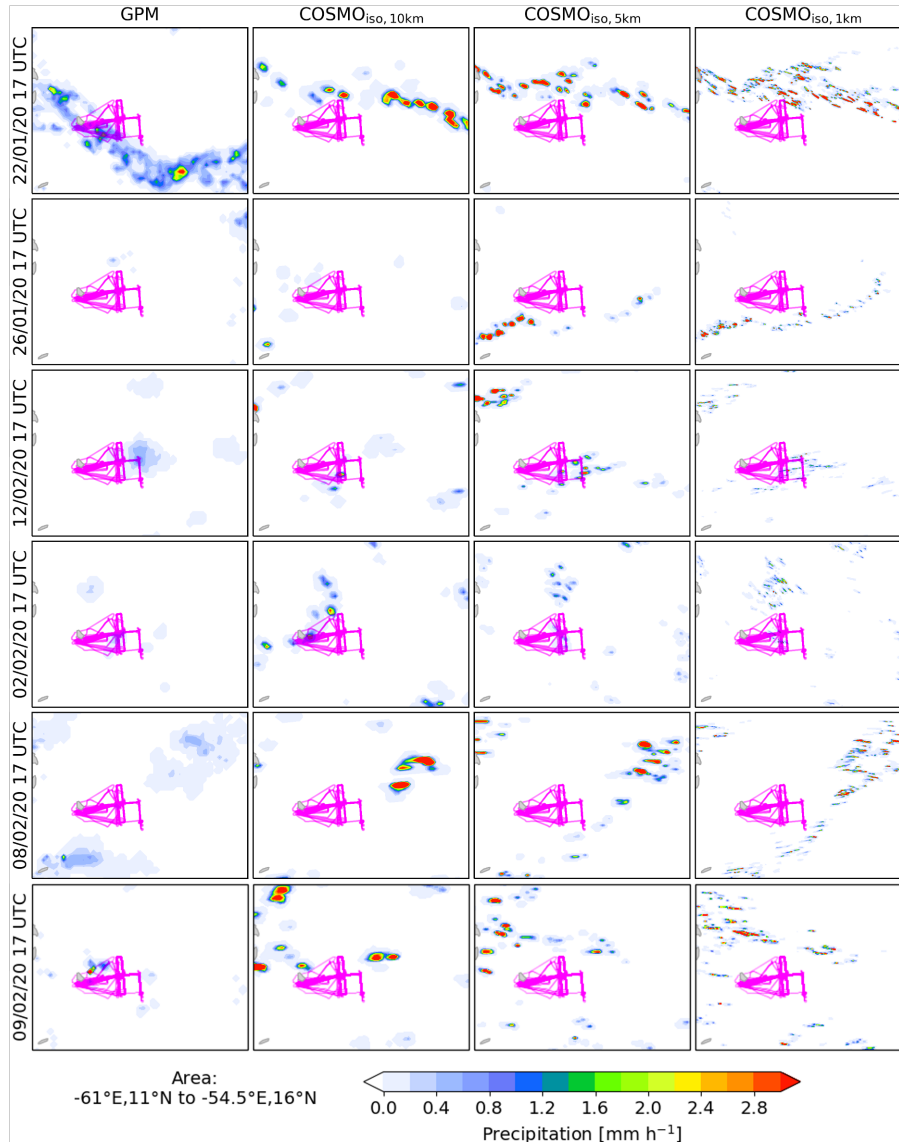
Figure 7. Hourly precipitation in the domain 54.5-61° W, 11-16° N from four datasets (columns) and on six dates (rows; also shown in Fig. 4). The first column shows the satellite-based precipitation estimate GPM (IMERG). The remaining columns show precipitation from the three COSMO$_{iso}$ simulations from 16:00-17:00 UTC.
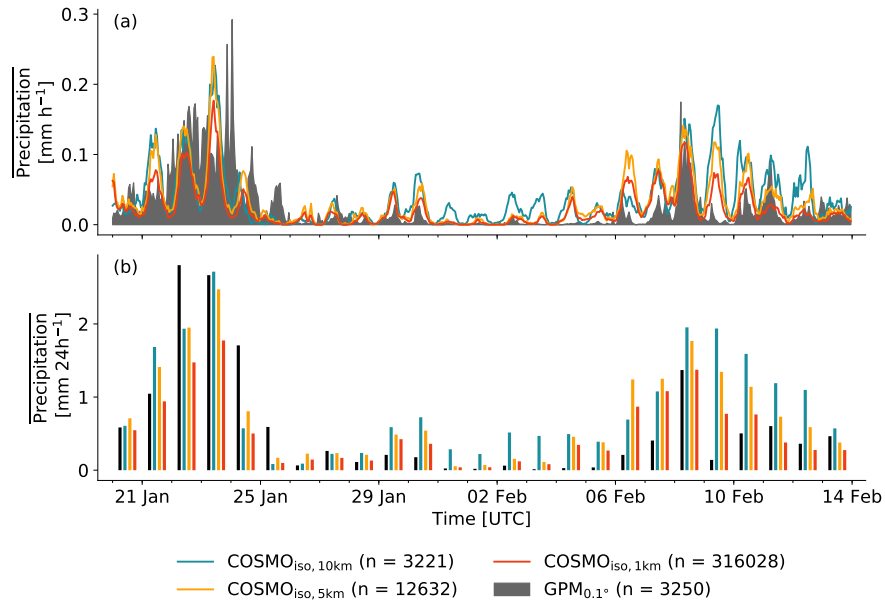
Figure 8. Time series displaying (a) hourly and (b) daily domain-averaged precipitation. Shown are the values from COSMO$_{iso,10km}$ (teal), COSMO$_{iso,5km}$ (yellow), COSMO$_{iso,1km}$ (red), and GPM (black). Only the data in the domain 54.5-61° W, 11-16° N are taken into account. The number of grid points in the domain is given in brackets.
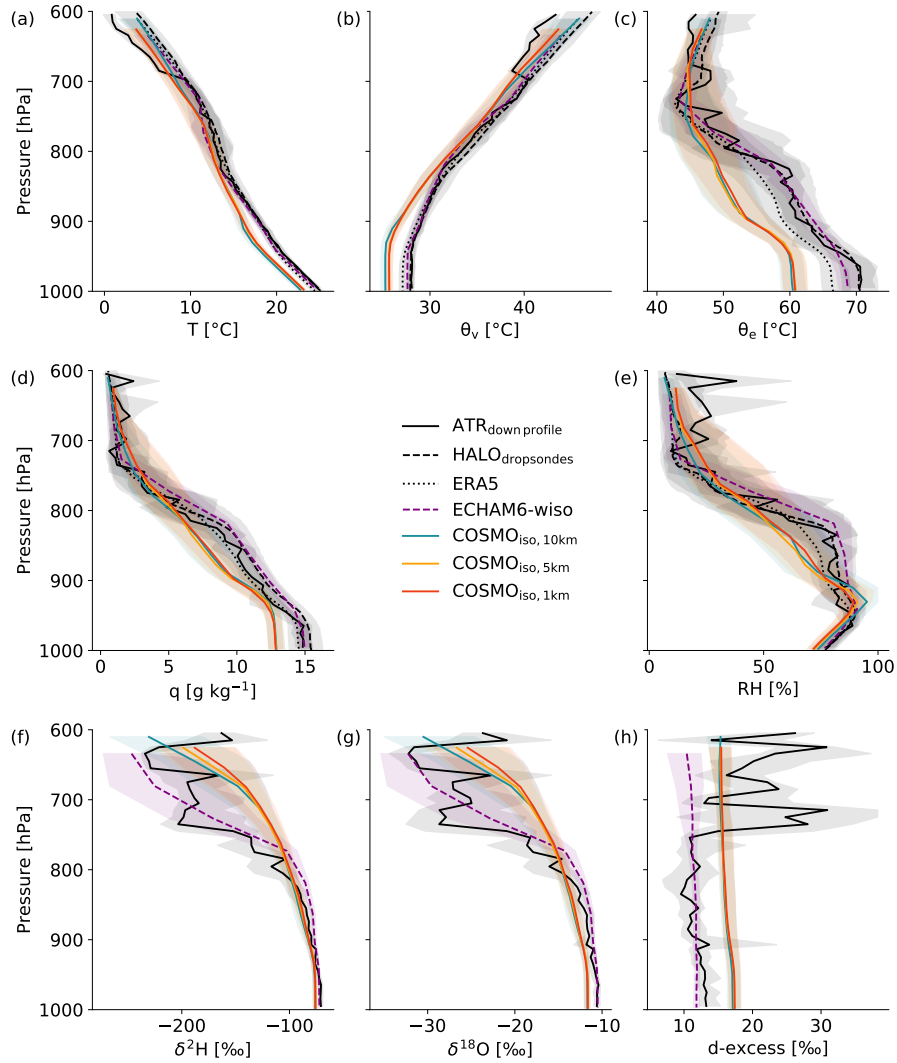
Figure 9. Vertical profiles of (a) temperature, (b) virtual potential temperature, (c) equivalent potential temperature, (d) specific humidity, (e) relative humidity, (f) δ^2H in vapour, (g) $\delta^{18}O$ in vapour, and (h) d-excess in vapour. Shown are the median (line) and the 25-75-percentile range (shading) of the ATR measurements (black continuous; only the data from the downward profiles from RF03-RF20), the HALO dropsondes (black dashed; 810 sondes), and the vertical profile closest to the centre of the EUREC^{4}A circle (at 57.717° W, 13.3° N) extracted at every hourly time step from 20 January to 13 February 2020 from ERA5 (black dotted; 600 profiles), ECHAM6-wiso (purple dashed; 100 profiles due to the 6-hh time steps), COSMO$_{iso,10km}$ (teal; 600 profiles), COSMO$_{iso,5km}$ (yellow; 600 profiles), COSMO$_{iso,1km}$ (red; 600 profiles). Similar profiles displaying the horizontal wind components are shown in Supplement 2.
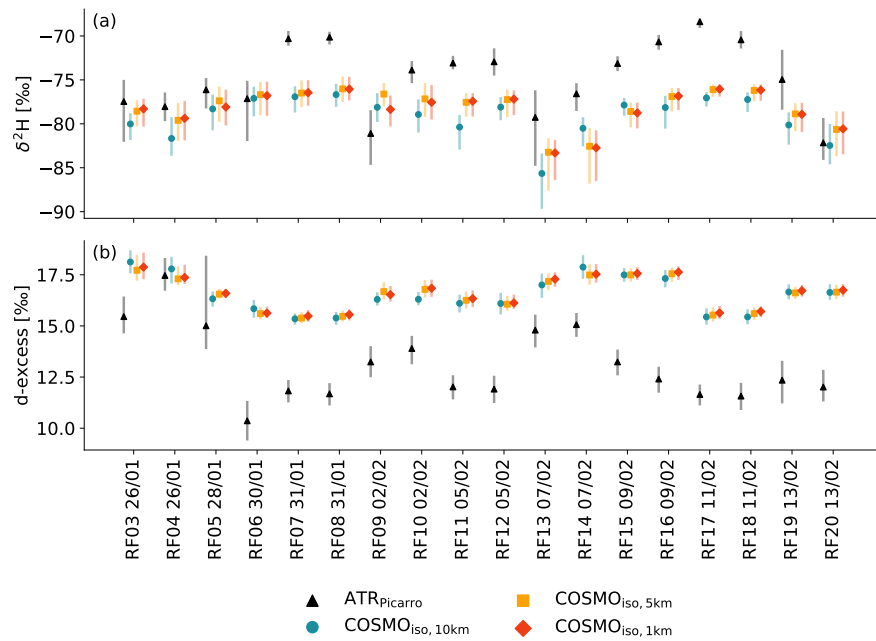
Figure 10. Median (marker) and the 25-75-percentile range (transparent line/errorbars) of cloud base (a) δ^2H, and (b) d-excess for each ATR flight. Shown are the observations from the ATR (black) and the simulations COSMO$_{iso,10km}$ (teal), COSMO$_{iso,5km}$ (yellow), and COSMO$_{iso,1km}$ (red). For the COSMO$_{iso}$ simulations, only the data from the hourly time steps that are closest to the ATR flights, and only the data in the domain 54.5-61° W, 11-16° N are taken into account.
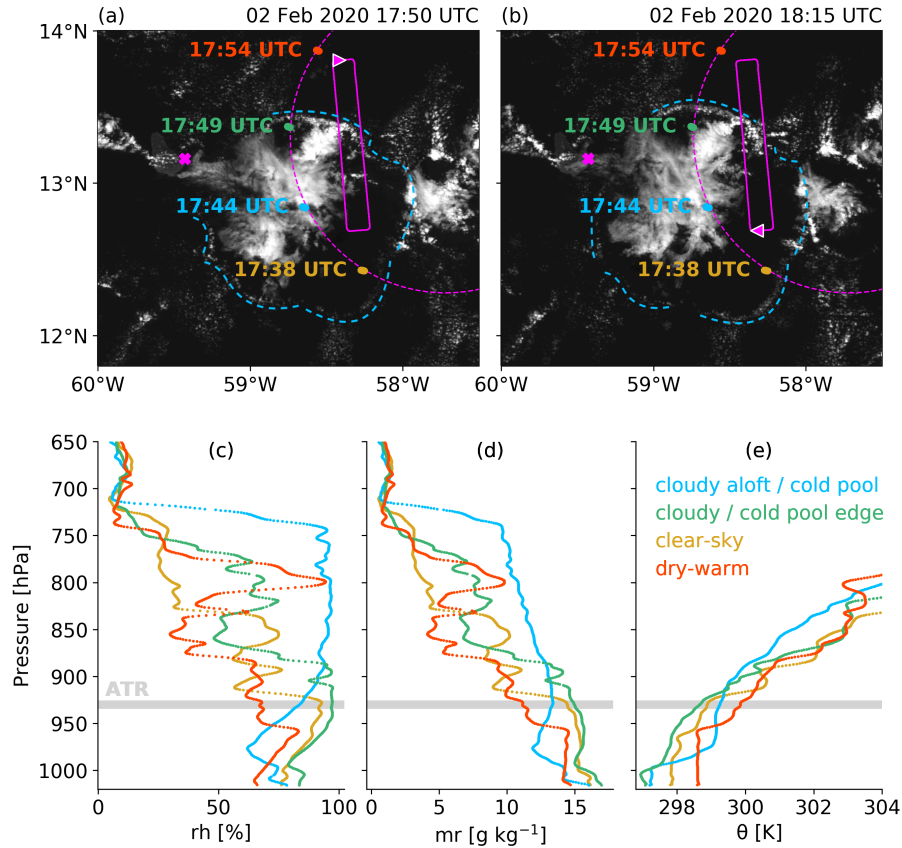
Figure 11. Cloudiness viewed by the visible channel of the GOES-16 satellite on 2 February 2020 together with the cloud base flight track of the ATR during RF10 (continuous pink) and the position of the ATR (pink triangle) at (a) 17:50 UTC and (b) 18:15 UTC (time steps chosen such, that the ATR is located at the northern and southern edge of the cloud base rectangle). The two maps also show the horizontal location of four dropsondes launched from the HALO during its circular flight pattern (dashed pink) with the launch time indicated to the left (yellow, blue, green, red). The blue dashed lines show the edges of the cold pool spreading radially away from the cloud in the centre of the satellite image. (c-e) Lower-tropospheric profiles of the four HALO dropsondes (shown in panel a,b) displaying (c) relative humidity, (d) mixing ratio, and (e) potential temperature. Note that the x-axis of (e) is set such that cloud base differences between the profiles are visible. The profile labels in (e) are intended to help the reader link the figure to the text.
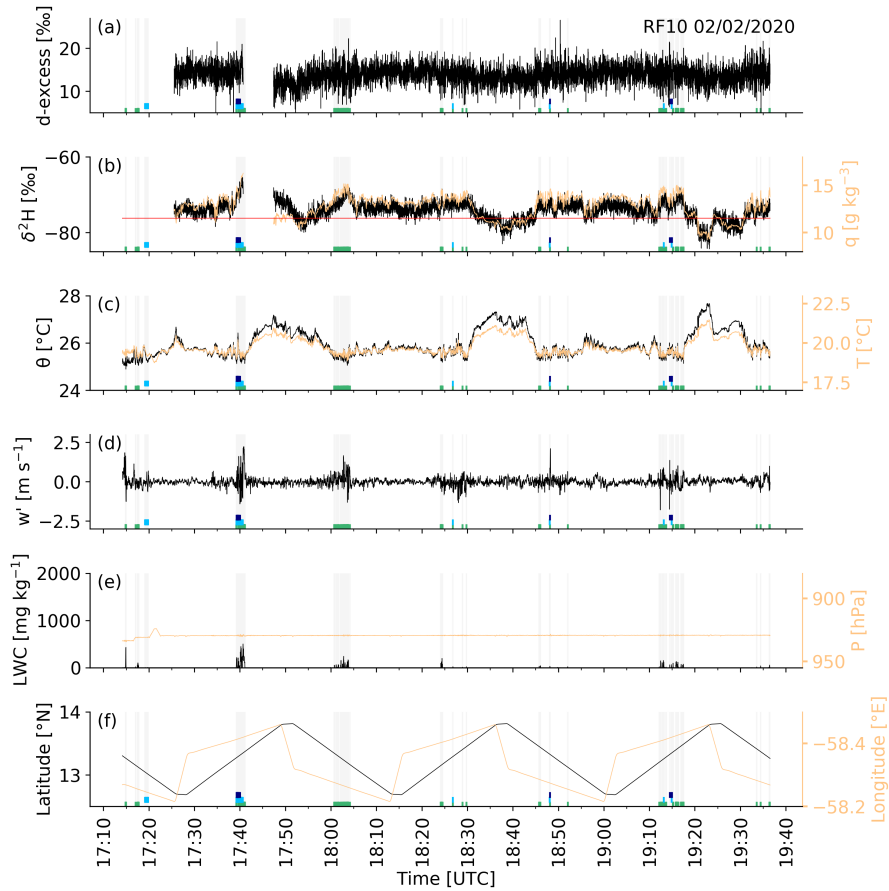
Figure 12. Time series of ATR measurements at cloud base during RF10 on 2 February 2020. Shown are (a) d-excess in vapour, (b) δ^2H in vapour (black) and specific humidity (sandy), (c) potential temperature (black) and temperature (sandy), (d) vertical wind velocity fluctuations relative to the flight-segment mean, (e) liquid water content (black) and pressure (sandy), (f) latitude (black) and longitude (sandy) of the ATR's position. In all panels, the grey shading indicates the presence of liquid water and the green/lightblue/darkblue bars at the bottom of (a,b,c,d,f) the cloud/drizzle/rain flags from the PMA dataset. Temperatures, vertical velocity, and positional data come from the ATR CORE dataset; d-excess, δ^2H, and specific humidity from the ATR Picarro dataset; and liquid water content from the ATR PMA dataset.
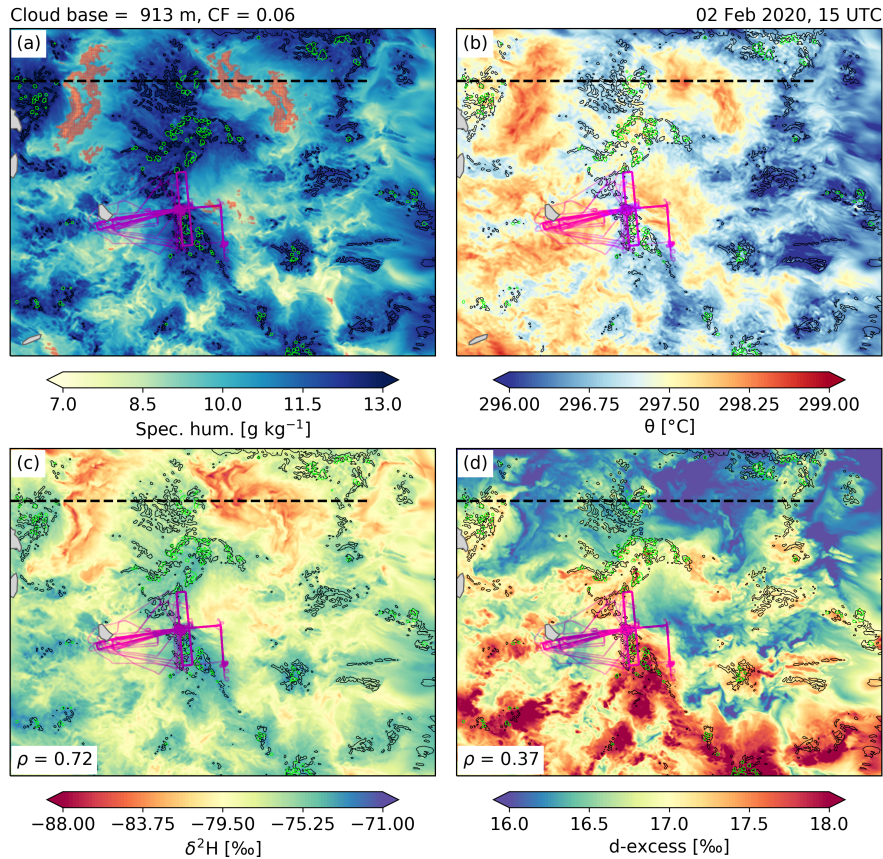
Figure 13. Fields of (a) specific humidity, (b) potential temperature, (c) δ^2H in vapour, and (d) d-excess in vapour at cloud base (z=913 m) in the domain 54.5-61° W, 11-16° N from the COSMO$_{\mathrm{iso,1km}}$ dataset at 15:00 UTC on 2 February 2020. The transparent red dots in (a) show grid points identified as *dry-warm* (see definition in text); the pink lines represent the track from all ATR flights; and the black line indicates the position of the cross-section shown in Fig. 14. Liquid cloud water contents of $10\,\mathrm{mg\,kg^{-1}}$ are shown as black contours and rain water contents of $1\,\mathrm{mg\,kg^{-1}}$ as green contours. In (c,d), the Pearson correlation coefficients (ρ) between the cloud base specific humidity and (c) δ^2H in vapour, (d) d-excess in vapour are indicated in the lower left corner. The cloud base cloud fraction (CF) at this time and in this region corresponds to 6 %.
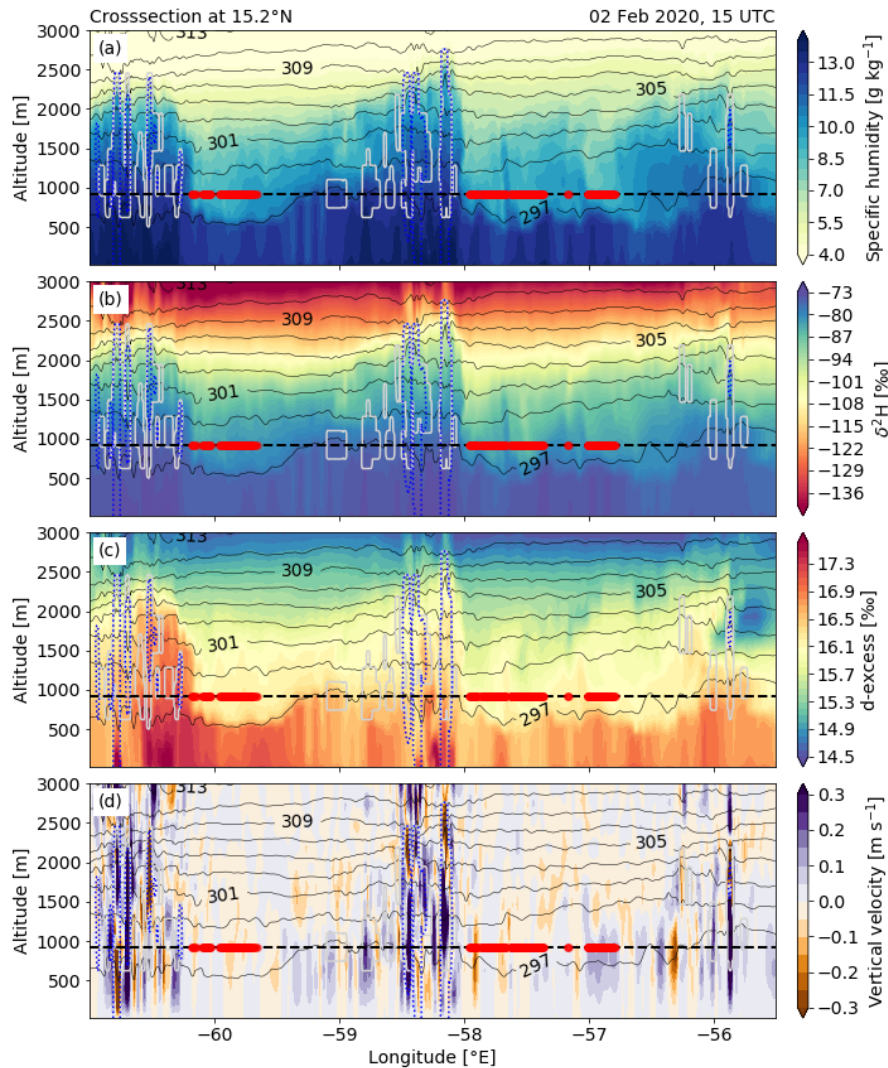
Figure 14. Cross-sections at 15.2° N of (a) specific humidity, (b) δ^2H in vapour, (c) d-excess in vapour, and (d) vertical velocity from COSMO$_{iso,1km}$ at 15:00 UTC on 2 February 2020. Shown is also potential temperature (black contours), cloud water contents of $10\,mg\,kg^{-1}$ (grey contours), rain water contents of $10\,mg\,kg^{-1}$ (dashed blue contours) and cloud base height (dashed black line; the level at which the fields are shown in Fig. 13) with the cloud base grid points identified as *dry-warm* (red dots; less than 0.01° latitude away from the cross-section). The location of the cross-section is indicated in Fig. 13.
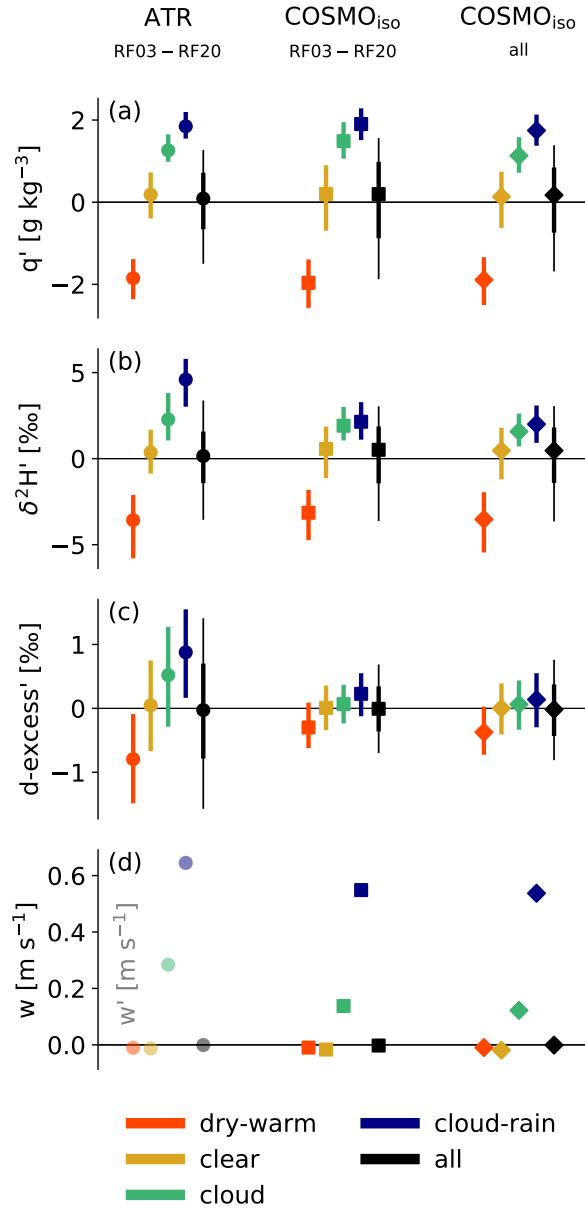
Figure 15. Distributions of the cloud base anomalies in (a) specific humidity, (b) δ^2H in vapour, and (c) d-excess in vapour of the data points identified as *dry-warm*, *clear*, *cloud*, *cloud-rain* (see the text for the definition) and all of them together (black). For the ATR data, anomalies are defined relative to each flight's mean cloud base conditions. For COSMO$_{iso,1km}$, anomalies are defined relative to each time step's mean cloud base conditions in the domain 54.5-61° W, 11-16° N. Shown are the median (markers), the 25-75-percentile range (lines), and for the overall distribution (black) the 10-90-percentile range (thin black line) of the cloud base data points collected by the ATR (RF03-RF20) and the COSMO$_{iso,1km}$ cloud base grid points at the hourly time steps that are closest to the ATR flights (COSMO$_{iso,RF03-RF20}$) and over the whole simulated period (20 January to 13 February 2020; COSMO$_{iso,all}$). (d) The mean absolute vertical velocity w for COSMO$_{iso}$ and mean vertical velocity fluctuations w' for the ATR (see text for the reasoning behind the variable choice) at cloud base. For the ATR, only RF10-RF19 provide vertical velocity measurements and are included here. Table 4 gives the percentage of data points in each category and Table 5 gives the numerical values of the median anomalies. Supplement 3 provides alternative versions of this figure.

Table 1. Model setup for the ECHAM6-wiso, COSMO$_{iso,10km}$, COSMO$_{iso,5km}$, and COSMO$_{iso,1km}$ simulations. The following abbreviations are used to describe the variables included in the nudging: surface pressure (SP), temperature (T), zonal wind component (U), and meridional wind component (V).

	ECHAM6-wiso	COSMO$_{iso,10km}$	COSMO$_{iso,5km}$	COSMO$_{iso,1km}$
Resolution	T127 (ca. 0.9°)	0.1° (ca. 10km)	0.05° (ca. 5km)	0.01° (ca. 1km)
Grid centre	-	45° W, 25° N	57.7° W, 13.3° N	57.7° W, 13.3° N
Grid points	$193 \times 139 \times 95$	$900 \times 800 \times 40$	$700 \times 600 \times 60$	$700 \times 600 \times 60$
Start	1 Dec 2019	6 Jan 2020	20 Jan 2020	20 Jan 2020
End	29 Feb 2020	13 Feb 2020	13 Feb 2020	13 Feb 2020
Time steps	6 h	1 h	1 h	1 h
Convection	Parameterised	Explicit	Explicit	Explicit
Boundary	ERA5	ECHAM6-wiso	COSMO$_{iso,10km}$	COSMO$_{iso,5km}$
Nudging	SP, T, U, V	U,V	U,V	U,V

Table 2. Root-mean-square differences of the median profiles (shown in Fig. 9) of the HALO dropsondes, ERA5, ECHAM6-wiso, COSMO$_{iso,10km}$, COSMO$_{iso,5km}$, and COSMO$_{iso,1km}$ data relative to the ones from the ATR$_{down\ profile}$ data over the layer 1000-850 hPa.

Dataset	T [°C]	θ_v [°C]	θ_e [°C]	q [g kg^{-1}]	RH [%]	δ^2H [‰]	δ^{18}O [‰]	d [‰]
HALO	0.1	0.1	0.8	0.8	2.8	-	-	-
ERA5	0.5	0.6	3.3	0.4	3.9	-	-	-
ECHAM6-wiso	0.4	0.3	1.6	0.8	4.7	3.1	0.4	1.0
COSMO$_{iso,10km}$	2.1	2.5	8.7	1.6	8.4	4.6	1.1	4.4
COSMO$_{iso,5km}$	1.9	2.2	8.6	1.7	8.7	3.9	1.0	4.6
COSMO$_{iso,1km}$	1.8	2.2	8.5	1.6	6.8	3.7	1.0	4.6

Table 3. Pearson correlation coefficients for $\delta^2 H$ in vapour and d-excess in vapour (d) between the median cloud base values from each ATR flight (RF03-RF20) and the hourly median values of all cloud base grid points in the domain 54.5-61° W, 11-16° N that are closest in time to the ATR flights from the three COSMO$_{iso}$ simulations.

Dataset	δ^2H [‰]	d [‰]
COSMO$_{iso,10km}$	0.65	0.74
COSMO$_{iso,5km}$	0.60	0.72
COSMO$_{iso,1km}$	0.69	0.72

Table 4. Percentage of data points assigned to one of the four cloud base categories (*dry-warm*, *clear*, *cloud*, *cloud-rain*). For the ATR, the time steps with valid isotope data are taken into account (RF03-RF20; note that these are fewer data points than available in the PMA dataset alone, because the isotope dataset has data gaps due to calibration or instrument failure, especially during rainy flight segments). For the $COSMO_{iso,1km}$, the data in the domain 54.5-61° W, 11-16° N is taken into account; once only from the hourly time steps that are closest to the ATR flights ($COSMO_{iso,RF03-RF20}$), and once for all simulated time steps ($COMSO_{iso,all}$).

Dataset	*dry-warm*	*clear*	*cloud*	*cloud-rain*
ATR	11.7 %	84.3 %	3.8 %	0.2 %
$COSMO_{iso,RF03-RF20}$	8 %	84 %	7.7 %	0.3 %
$COMSO_{iso,all}$	7.6 %	81.2 %	10.9 %	0.3 %

Table 5. Median values of the cloud base anomalies in specific humidity (q), δ^2H in vapour, and d-excess (d) in vapour for the data points identified as *dry-warm*, *clear*, *cloud*, *cloud-rain* (see the text for the definition) and the difference between the median of the cloud-rain and the dry-warm grid points (calculated with the exact values and not the rounded medians shown here) from the ATR and COSMO$_{iso,1km}$. For the ATR data, anomalies are defined relative to each flight's mean cloud base conditions. For COSMO$_{iso,1km}$, anomalies are defined relative to each time step's hour-of-the-day average cloud base conditions in the domain 54.5-61° W, 11-16° N. Shown are the median values over the cloud base time steps of the ATR (RF03-RF20), the COSMO$_{iso,1km}$ cloud base grid points at the hourly time steps that are closest to the ATR flights (COSMO$_{iso,RF03-RF20}$) and over the whole simulated period (20 January to 13 February 2020; COSMO$_{iso,all}$). Note that the 25-50-75 percentiles of these variables in each cloud base environment are shown in Fig. 15.

Dataset	Variable	*dry-warm*	*clear*	*cloud*	*cloud-rain*	*cloud-rain* minus *dry-warm*
ATR		−1.8	0.2	1.3	1.8	3.6
COSMO$_{iso,RF03-RF20}$	q [g kg^{-1}]	−2.0	0.2	1.5	1.9	3.9
COMSO$_{iso,all}$		−1.9	0.1	1.1	1.7	3.6
ATR		−3.6	0.4	2.3	4.6	8.2
COSMO$_{iso,RF03-RF20}$	δ^2H [‰]	−3.1	0.6	1.9	2.1	5.2
COMSO$_{iso,all}$		−3.5	0.5	1.6	2.0	5.5
ATR		−0.8	0.0	0.5	0.9	1.7
COSMO$_{iso,RF03-RF20}$	d [‰]	−0.3	0.0	0.1	0.2	0.5
COMSO$_{iso,all}$		−0.4	0.0	0.1	0.1	0.5