AGU ADVANCING EARTH AND SPACE SCIENCES

Earth and Space Science

RESEARCH ARTICLE
10.1029/2023EA003400

Evaluating TROPOMI δD Column Retrievals With In Situ Airborne Field Campaign Measurements Using Expanded Collocation Criterion

Iris Thurnherr[1,2], Harald Sodemann[1,3], Tim Trent[4,5], Martin Werner[6], and Hartmut Bösch[2,4,5]

[1]Geophysical Institute, University of Bergen, Bergen, Norway, [2]Institute for Environmental Physics, University of Bremen, Bremen, Germany, [3]Bjerknes Centre for Climate Research, Bergen, Norway, [4]Earth Observation Science, Department of Physics and Astronomy, University of Leicester, Leicester, UK, [5]National Centre for Earth Observation, Department of Physics and Astronomy, University of Leicester, Leicester, UK, [6]Alfred Wegener Institute (AWI), Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany

Key Points:

- New, expanded collocation criterion provides larger data sets for comparison of δD observations during a field campaign
- We evaluate TROPOMI δD column retrievals with in situ airborne measurements using a $COSMO_{iso}$-based collocation criterion
- The combination of subcolumn and total column averaged δD is used to characterize vertical δD gradients and reveal model biases

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

I. Thurnherr,
iris.thurnherr@env.ethz.ch

Citation:

Thurnherr, I., Sodemann, H., Trent, T., Werner, M., & Bösch, H. (2024). Evaluating TROPOMI δD column retrievals with in situ airborne field campaign measurements using expanded collocation criterion. *Earth and Space Science*, *11*, e2023EA003400. https://doi.org/10.1029/2023EA003400

Received 3 NOV 2023
Accepted 17 JUN 2024

Author Contributions:

Conceptualization: Iris Thurnherr, Harald Sodemann
Data curation: Tim Trent, Martin Werner
Formal analysis: Iris Thurnherr
Funding acquisition: Hartmut Bösch
Methodology: Iris Thurnherr, Tim Trent, Hartmut Bösch
Project administration: Hartmut Bösch
Supervision: Harald Sodemann
Visualization: Iris Thurnherr
Writing – original draft: Iris Thurnherr
Writing – review & editing: Harald Sodemann, Tim Trent, Martin Werner, Hartmut Bösch

Abstract

Satellite observations of column-averaged water isotopes are relatively new retrieval products that are in need of further in situ evaluation. Such evaluation studies are generally difficult to perform due to the wide mismatch in temporal and spatial scales between the satellite observations based on instantaneous pixel averages during an overpass and airborne in situ measurements ranging up to several hours over a km-scale. In addition, topography, weather conditions and in particular cloudiness impose severe constraints on an exact collocation between satellite and airborne in situ measurement platforms. Here we present a new method that allows a comparison between in situ measurements and satellite observations of δD on a broader statistical basis. We use regional isotope-enabled model simulations as intermediate information to identify the area for best comparisons. Applying our methodology to TROPOMI total column δD retrievals for the L-WAIVE campaign in Annecy, France, during June 2019 increases the number of satellite pixels for comparison despite widespread cloudiness on average by a factor of 20. In addition, the comparison of simulated and observed δD revealed a dependency of the satellite evaluation on the structure of the middle and upper troposphere. We conclude that our method provides a more robust statistic basis for in situ evaluation of δD satellite retrievals. The method will thus be useful in planning and executing forthcoming validation and evaluation campaigns, and can potentially be used for the evaluation of other satellite products.

Plain Language Summary

A characteristic of atmospheric water vapor is the concentration of stable heavy hydrogen and oxygen isotopes. Isotopic concentrations are observed by diverse techniques such as remote sensing from satellites and land-based instruments, and direct measurements from aircrafts, ships and stations on land. These measurements lead to a variety of data sets which span different distances (from a few to hundreds of km) and time periods (from seconds to days). While remote measurements provide data of a large spatial and temporal coverage at a coarse resolution, direct measurements are often obtained during research campaigns over a limited time period with high spatial and temporal resolution. Especially the later data sets are of high value as they describe the atmospheric state in high detail. In this study we develop a method to extrapolate direct, campaign based high-quality measurements of water vapor to the largest possible representative area using output from a numerical weather simulation. This allows us to compare remotely sensed and direct isotope measurements. This comparison illustrates how to interpret total column measurement and to identify where the model has difficulties to correctly simulate the vertical isotope distribution. This knowledge is of use for future application of remotely sensed data sets and model development.

1. Introduction

Clouds, precipitation, and radiative processes are intrinsically coupled to the addition, removal and transport of water vapor in the atmosphere. Phase changes, such as condensation and evaporation, create distinct signatures in the stable isotope composition of water vapor and precipitation (Gat, 1996). Therefore, stable water isotopologues (SWI) have long been recognized as a useful tool to trace the impact of different moist processes in the atmospheric water cycle. In recent years, space-based remote sensing of the deuterium isotopologue δD in water vapor has become possible from an increasing number of platforms (e.g., Boesch et al., 2013; Frankenberg et al., 2009; Herbin et al., 2007; Herman et al., 2020; Lacour et al., 2012; M. Schneider et al., 2016; Worden et al., 2007). Due to

their near-continuous, wide spatial coverage, such satellite observations of atmospheric δD are valuable for climatological and process studies of the atmospheric water cycle, and global-scale model validation. For example, total column and sub-column retrievals of δD have been used to assess the moisture turnover of the tropical troposphere (Noone, 2012; Worden et al., 2007), mixing of tropospheric air masses (Dahinden et al., 2021; Diekmann et al., 2021), or the evolution of tropical convective systems (Lacour et al., 2018; J. Lee et al., 2011).

A key requirement for such studies is a well-founded validation to ensure reliable data quality and accuracy. Satellite retrievals of δD have often been validated with ground-based remote sensing instruments, such as FTIR (Fourier-Transform Infrared Spectrometer) measurements (e.g., Frankenberg et al., 2013; M. Schneider et al., 2016; A. Schneider et al., 2020). Ground-based networks, such as the Total Carbon Column Observing Network (TCCON) and the Network for the Detection of Atmospheric Composition Change (NDACC), provide an essential backbone of observations that allow to tie satellite products to ground truth (Rokotyan et al., 2014; M. Schneider et al., 2006, 2012; Wunch et al., 2010). A subsequent evaluation of validated satellite products will then aim to use independent data sets to determine the usefulness and limits of interpretation for satellite remote sensing products.

In this context, in situ measurements of δD in water vapor are particularly valuable, as they provide detailed and direct spectroscopic evidence of SWI composition that can be used to evaluate satellite remote sensing products (Herman et al., 2014; M. Schneider et al., 2015). The availability of in situ data sets of water vapor isotope measurements has increased substantially in the last decade, including land-based (Aemisegger et al., 2012; Steen-Larsen et al., 2014; Weng et al., 2021), ship-based (Benetti et al., 2017; Bonne et al., 2019; Thurnherr et al., 2020) and aircraft-based measurements using commercially available laser spectrometers (Chazette et al., 2021; Henze et al., 2022; Herman et al., 2014; Salmon et al., 2019; Sodemann et al., 2017).

However, irrespective of the platform, such in situ δD measurements barely cover the spatial extent of typical satellite pixels, both vertically and horizontally. The challenges of spatio-temporal collocation of in situ and satellite observations is particularly obvious for airborne measurements. While an aircraft would take up to several hours to measure an atmospheric column, a satellite overpass would only last fractions of a second (Herman et al., 2014; M. Schneider et al., 2015). To extract the full value from in situ measurements for the evaluation of satellite products, specific approaches are needed to identify the time and space domain for which in situ measurements can be deemed representative based on plausible, physical principles.

Complex collocation definitions have been used previously for the satellite validation of trace gases such as CO_2 (Wunch et al., 2011). For example, a dynamic definition of collocation based on the correlation of potential temperature, total column CO_2 and transport modelling has been applied to compare CO_2 vertical profiles from satellite retrievals and aircraft measurements (Guerlet et al., 2013; Oshchepkov et al., 2012). There are important differences between CO_2 and SWI, such as the substantially shorter life cycle of water vapour (Gimeno et al., 2021; Sodemann, 2020), and the corresponding phase changes. Due to the connection of SWIs to the addition, transport and removal of water vapour from the atmosphere, SWI composition varies in response to atmospheric dynamics and thermodynamics. Thereby, short-term variations during the passage of weather systems can be as large as the meridional SWI gradients (Thurnherr et al., 2020).

Until now, only a constant time lag and distance criteria has been applied for the evaluation of δD data sets (e.g., Dahinden et al., 2021; Herman et al., 2014; M. Schneider et al., 2006; A. Schneider et al., 2020). Such simple criteria can be applied successfully if long time series are available, or in regions where cloud coverage is scarce, such as the subtropics, thus providing a sufficient number of collocation pixels. However, most research campaigns involving aircraft cover a limited time period and domain. Furthermore, in the mid-latitudes, satellite and airborne measurements are frequently compromised by cloud cover. As a result, only very few collocation pixels between satellite observations and in situ measurements may remain.

To substantially increase the number of collocated pixels for δD in a physically justified way, suitable collocation criteria need to consider the representativeness of in situ δD measurements. While air temperature can be a good proxy for meridional SWI gradients, synoptic scale variability is better represented by relative humidity and the air-sea temperature difference (Landwehr et al., 2021). Nonetheless, due to the complex interactions of processes determining SWI variability, it is difficult to approximate SWI variability using a single meteorological variable.

Global and regional isotope-enabled models contain the physical basis for the water vapour isotopic composition (e.g., Blossey et al., 2010; Brady et al., 2019; Cauquoin & Werner, 2021; Pfahl et al., 2012; Risi et al., 2010;

Yoshimura et al., 2008). Such models can provide access to the spatial and temporal representativeness of SWI, thus serving as an intermediary to define suitable collocation criteria. For a detailed overview of SWI modelling efforts see Risi et al. (2012), Galewsky et al. (2016), and Bailey et al. (2021). The isotope-enable regional numerical weather prediction model $COSMO_{iso}$ has been applied in various studies to investigate SWI variability on the time scales of weather systems to seasonal scales (e.g., Aemisegger et al., 2015; Dütsch et al., 2018; K.-O. Lee et al., 2019; Thurnherr et al., 2021). Similar to other regional isotope-enabled models, $COSMO_{iso}$ simulations provide physically plausible isotope information ranging from a few km to thousands of km, and from hours to weeks, thus covering the scales of both in situ measurements and δD satellite retrievals. Therefore, model simulations can be used to obtain an expanded collocation criteria, thus bridging the gap between in situ measurements and satellite remote sensing observations of δD on a wider scale.

Here we develop and apply an expanded collocation criterion to maximise the effectiveness of the evaluation of satellite retrievals of δD. Using in situ SWI data acquired during the L-WAIVE campaign (Chazette et al., 2021), we simulate weather situations over a European domain with the high-resolution regional isotope enabled model $COSMO_{iso}$. We compare δD in situ data with all TROPOspheric Monitoring Instrument (TROPOMI) satellite observations that are within a plausible range for comparison, based on the spatial variability of δD in the regional model. On this basis, we proceed to determine which key factors limit such a comparison, providing us with insight into both the representativeness of water isotope measurements under different weather situations, and the reliability of TROPOMI δD measurements. Based on this evaluation, we discuss under which conditions a refined approach to define collocation criteria optimised to SWI retrievals has the potential to widely increase the usefulness of in situ measurements for evaluation purposes.

2. Data Sets

This study uses two observational data sets: (a) in situ airborne measurements of δD during a field campaign in the French Alps and (b) satellite retrieved total column δD from TROPOMI. As an intermediary to bridge between the data sets, we use (c) $COSMO_{iso}$ model simulations that provide spatial and temporal continuity. In the following, these three data sets are described (Sections 2.1–2.3) and the approach to compare the three data sets is outlined (Section 2.4).

2.1. In Situ Measurements During the L-WAIVE Campaign

The Lacustrine-Water vApor Isotope inVentory Experiment (L-WAIVE) field campaign took place in June 2019 in the Annecy valley in the French Alps with the purpose to study the vertical distribution of water vapor over Lake Annecy (Chazette et al., 2021). During this campaign, a Picarro laser spectrometer L2130-i was installed on an ultralight aircraft to measure the isotopic composition of atmospheric water vapor during 11 flights that were distributed over 8 days, reached up to 3,500 m and had an effective temporal resolution of 10–30 s. These profiles will be used here to evaluate satellite retrieved δD.

The raw measurements of the isotope parameters were corrected for the water vapor mixing ratio - isotope composition dependency and normalized to the VSMOW-SLAP scale following routines developed at the FARLAB laboratory, University of Bergen, Norway (for details see Chazette et al., 2021, and references therein). The resulting vertical δD profiles are filtered to remove time periods of rapid elevation changes and periods of relative humidity above 90% indicating in-cloud measurements which might have led to condensation in the inlet line.

The synoptic weather situation plays an important role for the δD variability during the campaign and is therefore summarized here. The L-WAIVE campaign period was dominated by a high pressure system over Northern Africa, extending toward eastern Europe during the campaign (Chazette et al., 2021). This led to a strong daily cycle in the vertical profiles and a dominance of local convection with the development of convective precipitation cells in the vicinity of Annecy. A low pressure system passed north of the Alps on 15 June 2019 trailing a larger low over the British islands that prevailed from 12 to 16 June 2019. A detailed overview of the weather situation during the campaign is given by Chazette et al. (2021).

2.2. TROPOMI Total Column δD Retrieval

Measurements from the TROPOMI instrument onboard the Sentinel-5 Precursor (S5P) satellite are here used to retrieve total column averaged δD. TROPOMI is a hyperspectral nadir viewing imager comprising four grating spectrometers that measure Top Of Atmosphere (TOA) reflected and radiated solar radiation from the Earth's surface at ultraviolet/visible (UV/Vis), near-infrared (NIR) and shortwave infrared (SWIR) wavelengths. Operated since 13 October 2017, TROPOMI uses a push broom (non-scanning) configuration with a swath width of 2,600 km comprising 450 individual pixels, which produces typical pixel sizes (at nadir) of 7×28 km^2 in the UV, 7×3.5 km^2 for the Vis and NIR bands, and 7×7 km^2 for SWIR wavelengths (KNMI, 2017).

The University of Leicester (UoL) TROPOMI stable water vapor isotopologue (H2O-ISO) product has been produced as part of the S5P innovation (S5P-I) program run by the European Space Agency. The product is produced using the UoL full physics (UoL-FP) algorithm (Cogan et al., 2012), which has previously been used to retrieve water vapor and stable water vapor isotopologues from the Greenhouse gases Observing SATellite (GOSAT) (Boesch et al., 2013; Trent et al., 2018). For the generation of version 1 (V1.0.1) of the TROPOMI H2O-ISO product, the UoL-FP algorithm has been run in scaler profile mode with updates made to the a priori inputs, solar model and spectroscopy for SWIR bands, cloud screening, surface characterization, and uncertainty budget estimates. Further details can be found in the Algorithm Theoretical Basis Document (ATBD; Trent et al., 2021). The retrieval provides column-averaged estimates of the dry air mole fraction for H_2O and HDO (denoted by an X, i.e., XH_2O and XHDO, respectively) from which the total column-averaged δD (XδD) is calculated:

$$X\delta D = \left(\frac{\mathrm{XHDO}}{\mathrm{XH_2O}} \cdot (R_{\mathrm{VSMOW}})^{-1} - 1\right) \cdot 10^3, \tag{1}$$

where R_{VSMOW} is the Vienna Standard Mean Ocean Water (VSMOW) reference ratio equal to $3.11 \cdot 10^{-4}$. Initial comparisons of the TROPOMI H2O-ISO product against ground based in situ measurements from TCCON and MUSICA (MUlti-platform remote Sensing of Isotopologues for investigating the Cycle of Atmospheric water) NDAAC sites show the total column-averaged XδD to be negatively biased (M. Schneider et al., 2021). The $X\delta D$ bias depends on XH_2O and can be approximated with a linear least square fit. The following recommended bias correction from the S5P-I study is applied to each retrieved pixel: $X\delta D_{\mathrm{corrected}} = X\delta D_{\mathrm{retrieved,raw}} + (0.0112‰\ \mathrm{ppmv}^{-1} \cdot XH_2O - 1.03‰)$ with XH_2O of the corresponding pixel as retrieved with the UoL-FP algorithm from TROPOMI measurements. The mean bias correction for the data used in this study is +29.7‰ with a 25–75 percentile range of 22.8–36.0‰.

The data coverage from the satellite overpasses depends strongly on the cloud coverage on the corresponding dates. The S5P satellite overpasses Europe daily between 11:00 and 15:00. During the L-WAIVE campaign, there were six satellite overpass enabling total column averaged δD retrievals within 300 km from the campaign site from 12 to 22 June 2019. The extensive cloud coverage due to a low pressure system led to an absence of satellite retrievals on 14 and 15 June over the campaign region.

2.3. COSMO$_{\mathrm{iso}}$ Simulations During the L-WAIVE Campaign

COSMO$_{\mathrm{iso}}$ (Pfahl et al., 2012) is an isotope-enabled version of the limited-area regional numerical weather prediction model COSMO (Steppeler et al., 2003). The heavy water molecules $H_2^{18}O$ and $HD^{16}O$ are implemented in COSMO$_{\mathrm{iso}}$ with two additional water cycles analogous to and with the same physical processes as the $H_2^{16}O$ water cycle, except for isotopic fractionation during phase change processes. For a detailed description of the isotope parametrisations in COSMO$_{\mathrm{iso}}$ see Pfahl et al. (2012).

A 14-day COSMO$_{\mathrm{iso}}$ simulation starting at 0 UTC 10 June 2019 was conducted to cover the L-WAIVE campaign period from 12 to 22 June 2019. This simulation was performed with a horizontal resolution of 0.1° (approximately 10 km) on 40 vertical levels of which 19 levels are on average at a height below 3,500 m. Note, that the complex topography around Annecy is not resolved at a horizontal resolution of 0.1°. A sensitivity analysis with a higher horizontal resolution simulation of 0.02° (approximately 2 km) better represented the topography around Annecy but without an improvement of the model performance in this region (see Supporting Information S1). The model domain of the COSMO$_{\mathrm{iso}}$ simulation covers Western Europe and parts of the Atlantic Ocean,

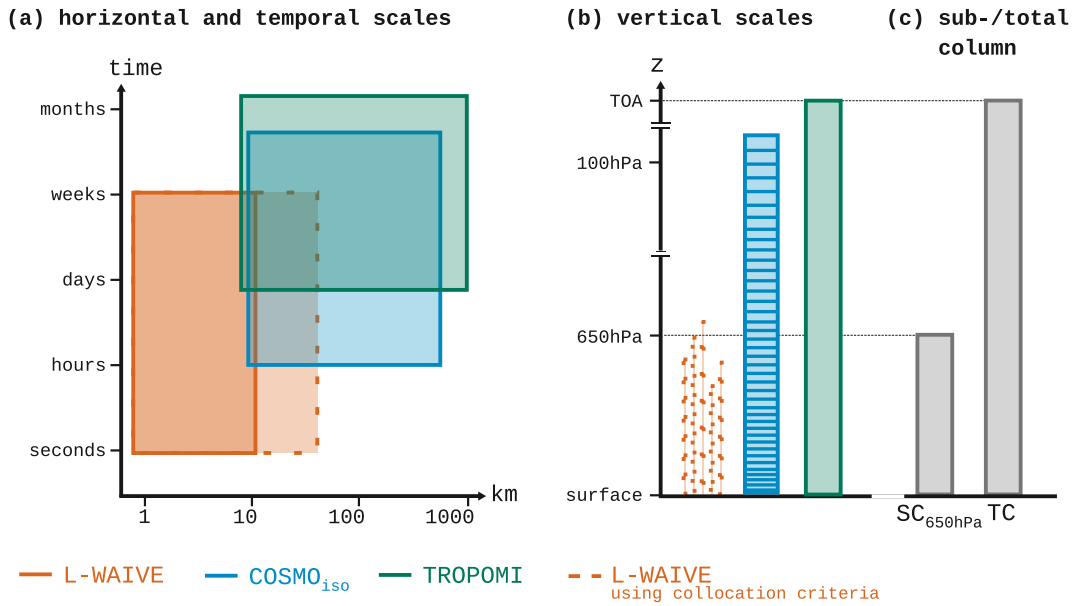

Figure 1. Overview of (a) horizontal and temporal, and (b) vertical scales of the in situ measurements during the L-WAIVE campaign (orange), the $COSMO_{iso}$ simulation (blue) and the TROPOMI retrieved δD (green). (c) Examples of a subcolumn up to a height of 650 hPa ($SC_{650\ hPa}$) and a total column (TC). TOA stands for Top Of the Atmosphere. The dotted orange line in (c) illustrates the L-WAIVE in situ measurements as points along a flight track.

Mediterranean, Baltic and Nordic Sea. A band of 50 km width along the model domain boundary has been removed to neglect effects from the boundary nudging. $COSMO_{iso}$ is coupled to the isotope-enabled land module TERRAiso which includes the parametrization of the prognostic isotopic composition of the soil, land-atmosphere intercept and snow. Isotopic fractionation during ocean evaporation is parametrized using the model by Craig and Gordon (1965) with a wind-speed-independent formulation of the non-equilibrium fractionation factor (Pfahl & Wernli, 2008).

The simulation was initialized and driven at the boundary with 6-hourly output from the global, isotope-enabled atmosphere model ECHAM6-wiso (Cauquoin & Werner, 2021) and spectrally nudged toward the ECHAM6-wiso wind fields above 850 hPa. Below 850 hPa, the $COSMO_{iso}$ simulation runs freely to minimize interference with the model's own representation of the water cycle. The ECHAM6-wiso simulation has a spatial resolution of 0.9°, 95 vertical levels and was spectrally nudged toward ERA5 reanalysis data (Hersbach et al., 2020) including surface pressure, vorticity, divergence and temperature.

Convection is treated explicitly in this simulation based on recent studies (e.g., Vergara-Temprado et al., 2019), which showed that convection parameterization schemes can be switched off in COSMO simulations at resolutions of the order of 10 km or higher to obtain a more realistic hydroclimate over Europe. A comparison of $COSMO_{iso}$ simulations at a horizontal resolution of 7 and 14 km with and without convection parametrization showed a good agreement between the simulations and with satellite observations of the isotopic composition of water vapor over the West African monsoon (de Vries et al., 2022). These results imply that explicit convection can adequately represent the atmospheric water cycle at a horizontal model resolution of 10 km.

2.4. Spatial and Temporal Coverage of Data Sets

In this study, we develop a method to evaluate satellite retrieved δD with in situ measurements of δD in water vapor. The main challenge of this comparison is the different temporal and spatial coverage of the two data sets (Figure 1a). The L-WAIVE measurements cover a time span of seconds to a week and a horizontal spatial range of a few meters to 10 km. The TROPOMI data set spans a horizontal distance of 7 km to several 1,000 km on a time span of hours to years. As an intermediary between these two data sets, the $COSMO_{iso}$ simulation spans a time period of 1 hr to 2 weeks and 10–800 km.

As illustrated in Figure 1b, the data sets do not only differ in their horizontal coverage but also in their vertical coverage and resolution. The L-WAIVE measurements have a high vertical resolution up to a few meters and cover a vertical column up to 3,500 m. The $COSMO_{iso}$ simulation has up to 10 m vertical resolution that is decreasing with increasing height and covering the vertical column up to 4 hPa with 19 (of total 40) vertical model levels below 3,500 m. The TROPOMI retrieval instead provides only total column values, thus has the lowest vertical resolution of the three data sets but covers the largest vertical extent. To be able to compare the data sets, the L-WAIVE and $COSMO_{iso}$ δD have to be translated into column averages, either for a subcolumn representative for part of the column up to a specific height calculated with the in situ L-WAIVE measurements (up to 3,500 m) and $COSMO_{iso}$ (for 16 pressure levels) or the total column calculated with $COSMO_{iso}$ (see Figure 1c). To improve readability we will refer to total column δD as *TC* and subcolumn δD as *SC* in the remainder of the manuscript. The calculations of SC and TC are outlined in the next section (Section 3.1).

3. Method

The in situ measurements, TROPOMI retrievals and the $COSMO_{iso}$ simulation need to be translated into a space where the data sets are comparable. For this purpose, the in situ measurements and the $COSMO_{iso}$ simulation are translated into column averaged values as described in Section 3.1. The TROPOMI retrievals are selected within a plausible region around the location of the in situ measurement. This region is defined by a collocation mask as introduced in Section 3.2.

3.1. Calculation of Total and Subcolumn-Averaged δD

To be able to compare the data sets, the L-WAIVE and $COSMO_{iso}$ vertical scales are translated into column averages (SC and TC) based on the pressure weighting function h' (O'Dell et al., 2012, see also Appendix A). First, the total column vapor mixing ratio is calculated for H_2O and HDO.

$$XH_2O = \sum_{i=1}^{z'} h'_i m'_{H2O,i}, \tag{2}$$

where z' is the number of vertical layers, h'_i is the pressure weighting function, which is taking into account the mass of air in each vertical layer, and $m'_{H2O,i}$ is the mean layer water vapor mixing ratio. XHDO is calculated accordingly with m'_{HDO}:

$$XHDO = \sum_{i=1}^{z'} h_i m_{HDO,i}. \tag{3}$$

Second, the column averaged δD is then defined as

$$\delta^2H = \left(\frac{R_{COL}}{R_{VSMOW}} - 1\right) \cdot 1000[], \tag{4}$$

where $R_{COL} = \frac{X_{HDO}}{X_{H2O}}$ is the column averaged isotopic ratio for deuterium and $R_{VSMOW} = 3.11 \cdot 10^{-4}$ the reference isotopic ratio for deuterium of the Vienna Standard Mean Ocean Water (VSMOW).

Analogous to column-average XH_2O, column-averaged specific humidity is calculated based on h' and will be referred to as *Xq* in the following.

Depending on the data set, either TC, using all vertical layers to the models upper bound at 4 hPa, or SC, spanning only a sub-column from the surface to a specific height, or both, is calculated. Due to the limited vertical coverage of the in L-WAIVE situ data, SC is calculated up to the respective flight height of the L-WAIVE in situ profiles. The SC is calculated by binning the vertical profiles to pressure layers and applying the pressure weighting function based on the humidity, temperature and pressure of these layers. For a comparison of the simulated δD distribution in $COSMO_{iso}$ with the TROPOMI TC retrievals, the TC was calculated from the $COSMO_{iso}$ output fields. For each grid point, the $COSMO_{iso}$ output fields were interpolated on 16 regularly spaced pressure levels between the surface pressure and 80 hPa. These levels mimic the pressure levels used for the satellite retrievals.

Based on these 16 levels, TC is calculated using Equations 2–4. In addition, SC is calculated for the subcolumn underneath each of the 16 pressure levels in the $COSMO_{iso}$ simulation.

3.2. Collocation Regions as a Measure of Spatial Representativeness

For the evaluation of satellite retrievals with in situ δD measurements, we need to define collocation regions as a measure of the spatial representativeness of the in situ point measurements. The term *spatial representativeness* refers here to the wider region for which a point measurement is representative in terms of showing a similar temporal evolution of δD. We define collocation regions using simulated TC and two statistical measures: (a) the Pearson correlation coefficient (ρ) and (b) the root mean square deviation (RMSD). ρ is used to study the correlation of TC in time between two points in the $COSMO_{iso}$ domain. RMSD will give information on the offset of TC in time between two grid points. The combination of these two measures enables to take relative and absolute spatial changes in δD into account in defining spatial representativeness.

The collocation mask is defined by thresholds of ρ and RMSD. To consider a location to represent the same air masses as the measurement location, a high ρ and low RMSD of TC over a given time period at this location with respect to the measurement location is required. Therefore, we aim to find a threshold value for ρ and RMSD that satisfies these conditions. These collocation mask thresholds can be defined in at least two ways: a constant value can be defined for both statistical measures or a time-dependent threshold can be applied, thereby including the temporal TC variability in the region of interest. Based on the high temporal variability of SWIs a time-variant threshold is used here to best represent the TC variability simulated around Annecy.

We define the $\rho_f(t)$ and $RMSD_f(t)$ threshold for each 1-hourly model output time step t in the following way: ρ and RMSD of TC at Annecy with TC of every model grid point for a time window of $t \pm 24$ hr are calculated. The collocation mask thresholds $\rho_f(t)$ and $RMSD_f(t)$ are then defined in the following way:

$$\rho_f(t) = \min(\rho(t)) \quad \text{with } \rho(t) > 0 \text{ and } \overline{\mathrm{RMSD}(t, A_{\rho(t)})} < f \cdot \overline{\mathrm{RMSD}(t)}, \tag{5}$$

where $\overline{\mathrm{RMSD}(t)}$ is the mean RMSD over the full model domain at time t, $\overline{\mathrm{RMSD}(t, A_{\rho(t)})}$ is the mean RMSD at time t within $A_{\rho(t)}$, which is the area encircled by a single closed contour of the value ρ around Annecy, and $0 \leq f \leq 1$. Based on the definition of $\rho_f(t)$,

$$\mathrm{RMSD}_f(t) = \overline{\mathrm{RMSD}(t, A_{\rho_f(t)})}. \tag{6}$$

The resulting thresholds represent a region with a distinctly lower RMSD than the full model domain. The mean RMSD within the collocation region is less than a fraction f of the RMSD over all model grid points with an as high as possible ρ (see Figure 2 illustrating the calculation of ρ_f and $RMSD_f$ for 12 UTC 13 June 2019). The fraction f is not defined a-priori and has to be chosen individually based on the study area. This collocation method will be applied to the L-WAIVE campaign period in Section 5 to demonstrate the gain in data points compared to previously used collocation regions.

4. Comparability of $COSMO_{iso}$ Model and SWI Observations

To use $COSMO_{iso}$ as an intermediary between the in situ and the TC retrievals, we need to validate the model to ensure that it adequately represents the δD variability on the scales of the observational data sets. We will first compare $COSMO_{iso}$ vertical interpolated profiles to the L-WAIVE in situ measurements. Secondly, we will validate $COSMO_{iso}$ TC with satellite retrieved TC.

4.1. Comparison of In Situ and $COSMO_{iso}$ Vertical δD Profiles

To assess how well $COSMO_{iso}$ simulates δD in the Annecy valley, vertically binned δD profiles of the in situ L-WAIVE measurements and the $COSMO_{iso}$ output linearly interpolated along the L-WAIVE flight tracks are compared. The vertically binned in situ δD profiles from the L-WAIVE campaign show relatively weak vertical gradients below 800 hPa (Figure 3a). Above 800 hPa, the profiles show a decrease in δD with height starting at varying heights and with different slopes between the flights. Exceptions from weak vertical gradients up to

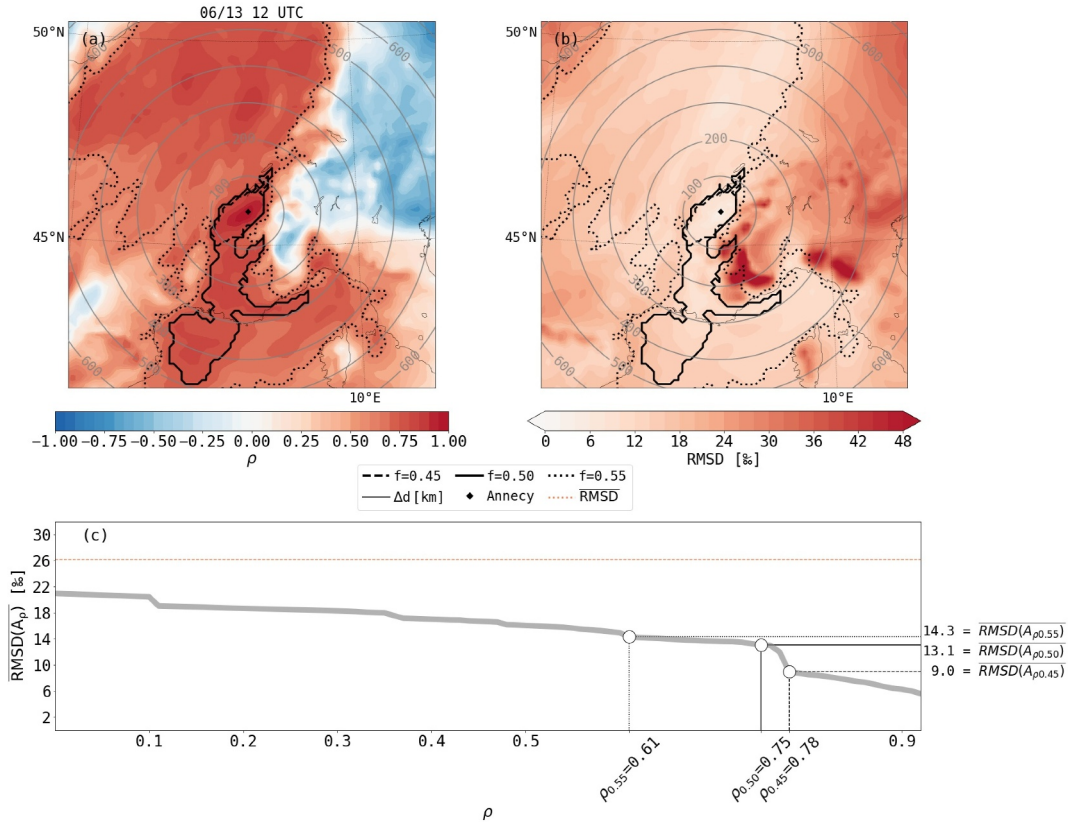

Figure 2. Pearson's correlation coefficient ρ (a) and root mean square deviation RMSD (b) between TC at 12 UTC 12/06/2019 and 12 UTC 14/06/2019 at Annecy with any grid point in the $COSMO_{iso}$ domain is shown. The black contours show the collocation masks for $f = 0.45$ (dashed line), 0.50 (solid line) and 0.55 (dotted line), and the gray contour the distance in 100 km steps from Annecy (black diamond). Panel (c) shows the mean RMSD for all grid points within a closed ρ contour around Annecy ($\overline{RMSD(A_\rho)}$, gray line) as a function of ρ. The three markers show the thresholds $\overline{RMSD(A_{\rho f})}$ and ρ_f for the masks of $f = 0.45$ (dashed line), 0.50 (solid line) and 0.55 (dotted line) as shown in panels (a) and (b).

800 hPa are the profiles on 16 and 17 June 2019 (light blue lines) when a strong moisture gradient was observed around 800 hPa (Figure 3b) as also seen from 16 to 17 June in ground-based lidar measurements (Chazette et al., 2021). The mean SC over all flights is −112.4‰.

The simulated δD profiles (Figure 3b) show qualitatively similar vertical gradients as the measurements. Below 800 hPa δD varies between −110 and −80‰. Above 800 hPa, there is often a weaker δD decrease with height in the simulation than observed in the measurements. Nonetheless, the correlation between δD along the flight track in the in situ measurements and the $COSMO_{iso}$ simulation is above 0.6 for most of the flights (see ρ in Table 1). This indicates that the δD variability along the flight tracks is well captured by $COSMO_{iso}$.

The difference between the measured and simulated δD profiles (Figure 3c) shows a relatively constant positive bias between 0 and 25‰ for each flight profile in $COSMO_{iso}$ below 800 hPa and strongly varying biases above 800 hPa ranging from −55 to 200‰. The mean bias in the modeled SC δD over all flights is 11.1‰. Even though this bias seems large compared to δD variability at the surface, the bias is an order of magnitude smaller than vertical δD gradients and is smaller than the bias in the satellite measurements relative to ground-based remote sensing. The positive bias in modeled SC agrees with too high simulated specific humidity q compared to the measurements with a mean subcolumn bias in averaged q of 1.0 g kg^{-1} (Figure 3d). While δD and q biases agree in sign and magnitude, air temperature shows a mean subcolumn bias of −2.2°C (not shown). Consequently, relative humidity in the lower troposphere is higher in $COSMO_{iso}$ than the measurements. The spread of the eight closest model grid points around the interpolated flight track (error bars in Figure 3b) indicate strong vertical δD gradients related to the position of the boundary layer height in $COSMO_{iso}$. A displacement in the modeled

10.1029/2023EA003400

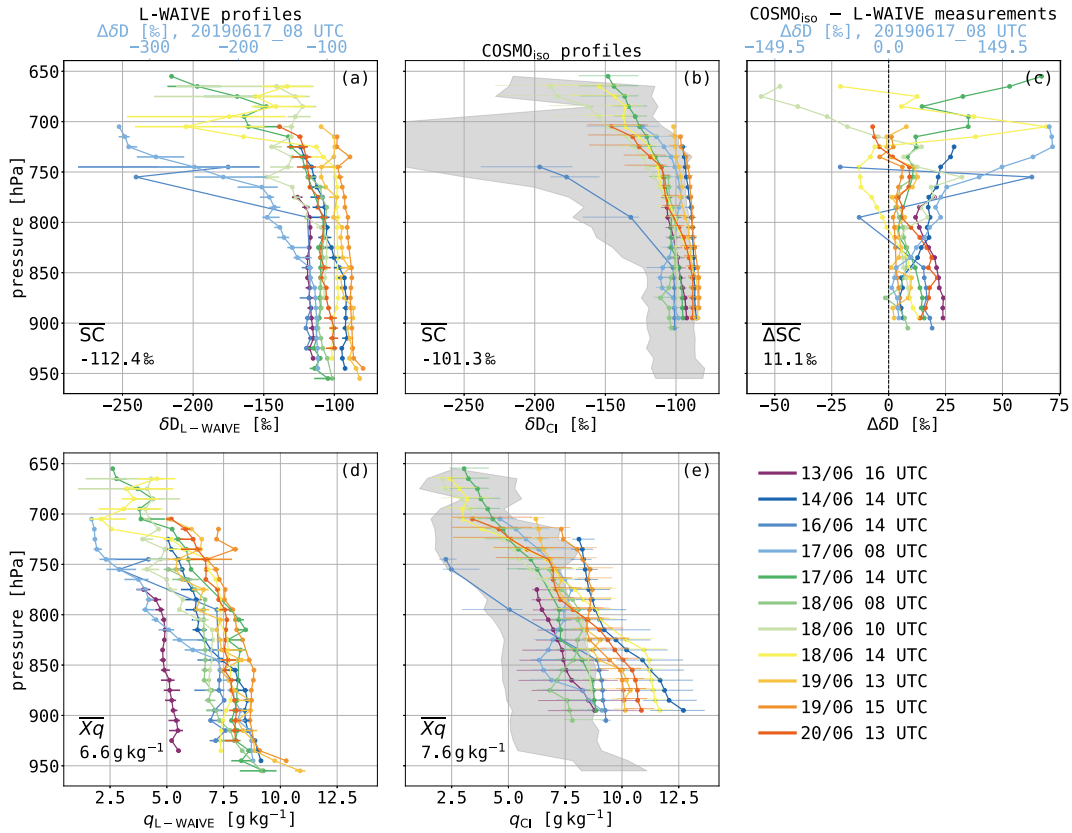

Figure 3. Vertical profiles of δD (first row) and specific humidity q (second row) binned into pressure bins for in situ measurements from the L-WAIVE campaign (a, d) and $COSMO_{iso}$ subcolumns linearly interpolated along the L-WAIVE flight track (b, e). Each profile represents a different fligth (in colors). The horizontal lines in panels (a, d) show the 5–95 percentile range of the 30 s resolution measurements and in panels (b, e) the binned minimum to maximum values of the eight closest gridpoints in $COSMO_{iso}$ around the flight path. The gray shaded area in panels (b, e) denote the range of the measured profiles (as shown in panels (a, d), respectively). Panel (c) shows the difference between the measured and simulated δD profiles. Please note that the x-axis for the flight at 8 UTC 17/06/2019 (light blue line) is shown at the top in panel (a and c) (light blue x-labels) and covers a different range than the other flights. The mean over all vertical profiles of sub-column averaged δD ($\overline{SC}$ in a, b), $\Delta\delta$D ($\overline{\Delta SC}$ in c) and q ($\overline{Xq}$ in d, e) is shown in the lower left corner of the respective panel.

boundary layer top or a different modeled δD gradient across the boundary layer top can lead to relatively large biases at this height.

We conducted a series of sensitivity experiments with $COSMO_{iso}$ to test possible reasons for the observed model-measurement differences (see Supporting Information S1 for details). Neither increasing the horizontal model resolution to 0.02° nor changing the land-atmosphere isotope parametrization in $COSMO_{iso}$ had a significant impact on the reduction of the δD bias in the boundary layer. The ECHAM6-wiso initial and boundary data, which shows a negative bias in boundary layer δD, can also not explain the positive δD bias. Possible origins of this bias are further discussed in Section 7. Even though a positive δD bias in $COSMO_{iso}$ will also be translated into SC, the relative temporal changes in SC can still be used to understand the effect of vertical column averaging for a comparison with satellite data.

From this comparison of measured and simulated vertical δD profiles, we conclude that there is a qualitatively good agreement between observations and simulation. Even though the decrease in δD from the boundary layer toward the free troposphere is often too weak in $COSMO_{iso}$, the temporal and spatial variability in vertical δD gradient is represented in the $COSMO_{iso}$ simulation.

Table 1
Difference in SC (ΔSC) and Pearson Correlation of δD (ρ) Between $COSMO_{iso}$ and L-WAIVE Measurements for Each Flight and Difference in TC (ΔTC) Between $COSMO_{iso}$ and TROPOMI Retrieval for Each Satellite Overpass During the L-WAIVE Campaign

Flight	ΔSC [‰]	ρ
16 UTC 13/06	18.7	0.46
14 UTC 14/06	11.1	0.93
14 UTC 16/06	12.2	0.79
08 UTC 17/06	36.4	0.66
14 UTC 17/06	10.8	0.92
08 UTC 18/06	5.2	0.48
10 UTC 18/06	7.2	0.44
14 UTC 16/06	12.2	0.79
14 UTC 18/06	4.3	0.72
13 UTC 19/06	2.9	0.76
15 UTC 19/06	1.8	0.76
13 UTC 20/06	11.5	0.90
Mean	11.1	0.72
Overpass	**ΔTC [‰]**	**Pixels**
12 UTC 12/06	−51.8 [−83.0 to −11.9]	690
12 UTC 13/06	−11.7 [−41.6 to 17.7]	1,562
13 UTC 15/06	25.5 [14.7 to 37.6]	854
13 UTC 16/06	−0.6 [−10.6 to 17.3]	930
12 UTC 17/06	12.7 [−1.1 to 30.5]	1,341
12 UTC 22/06	−27.6 [−56.2 to −5.0]	244
All overpasses	−4.1 [−27.2 to 24.6]	5,621

Note. The values in brackets denote the 25 to 75 percentile range. The last column shows the number of pixels available for the comparison of each and all overpasses.

4.2. Comparison of TROPOMI Retrieved and $COSMO_{iso}$ TC

As a second step of assessing the δD variability in $COSMO_{iso}$, we now compare the retrieved TROPOMI TC with TC from the model simulation. As an example, Figures 4a and 4b shows TC for 13 June 2019 between 11:55 and 12:01 UTC for a satellite overpass and at 12 UTC in the $COSMO_{iso}$ simulation. Due to extended cloud cover, retrieved TCs are only available for regions to the northwest of the Alps, most of Italy, and continental Eastern Europe (Figure 4a).

Cloud covered areas in $COSMO_{iso}$ correspond mostly to missing values over the continent in the satellite retrieval indicating that $COSMO_{iso}$ catches the overall cloud patterns over Europe. Only pixels that are cloud free in both the satellite observations and $COSMO_{iso}$ are used in the following. Generally, $COSMO_{iso}$ is able to reproduce the large-scale TC pattern for this overpass with low TC over the Alps, to the northwest of the Alpine ridge and along the western Adriatic coast, and high TC over Eastern Europe (Figure 4b). The difference between the two data sets shows a high spatial variability (Figures 4c and 4d). While $COSMO_{iso}$ shows a positive bias around the Alps where TC was influenced by a cyclone passage, there is a strong negative bias over Eastern Europe dominated by a high pressure system.

For this overpass, the mean difference in TC between $COSMO_{iso}$ and the satellite retrieval is −11.7‰ [−41.6 to 17.7‰] (brackets denote 25–75 percentile range). Most of the satellite overpasses during the L-WAIVE campaign show a negative mean TC bias in $COSMO_{iso}$ compared to the satellite retrieved TC for the full domain (see Table 1). Over all overpasses, the difference in TC between $COSMO_{iso}$ and TROPOMI is −4.1‰. In contrast to δD, total column averaged specific humidity Xq in $COSMO_{iso}$ shows a positive mean bias of 0.26 g kg^{-1} [0.05–0.46 g kg^{-1}]. This bias is similar to the $COSMO_{iso}$ q bias with respect to the in situ measurements in the lower troposphere. The positive q bias in the boundary layer is thus also seen in the total column, while the positive δD bias in the boundary layer is not seen in TC.

Even though there is an absolute difference between $COSMO_{iso}$ and the observations as shown in Table 1, the spatial and temporal variability is represented well in the $COSMO_{iso}$ simulation (Figure 4; see also Figures S1 and S2 in Supporting Information S1). This allows to study the horizontal and temporal TC variability in the $COSMO_{iso}$ simulations as a measure of the spatial scale of δD patterns in atmospheric water vapor using suitable collocation criteria.

5. Evaluation of TC Retrievals With In Situ Measurements

We will now apply the introduced methodology to our observational data sets during the L-WAIVE campaign to illustrate in a case study how an evaluation of satellite retrievals with in situ measurements can be done. First, the collocation masks for the L-WAIVE campaign period are generated. Second, the collocation masks are used to identify satellite pixels for the evaluation with in situ measurements at Annecy.

5.1. Collocation Mask for L-WAIVE Case Study

To generate the collocation mask for the L-WAIVE campaign period, the factor f (see Section 3.2) has to be defined. Figure 5 shows the collocation masks at Annecy for each satellite overpass of this study using three different f. For the first three overpasses, f has a large influence on the mask area and shape, while for the last three overpasses different f create mostly the same mask. All masks have in common that they follow the Alpine ridge and exclude regions to the southeast of the Alps. This represents the role of the Alps as a moisture barrier leading to different moisture histories on either side of the mountains and, thus, weakly correlated δD in water vapor (see low ρ to the East of the Alpine ridge in Figure 2a).

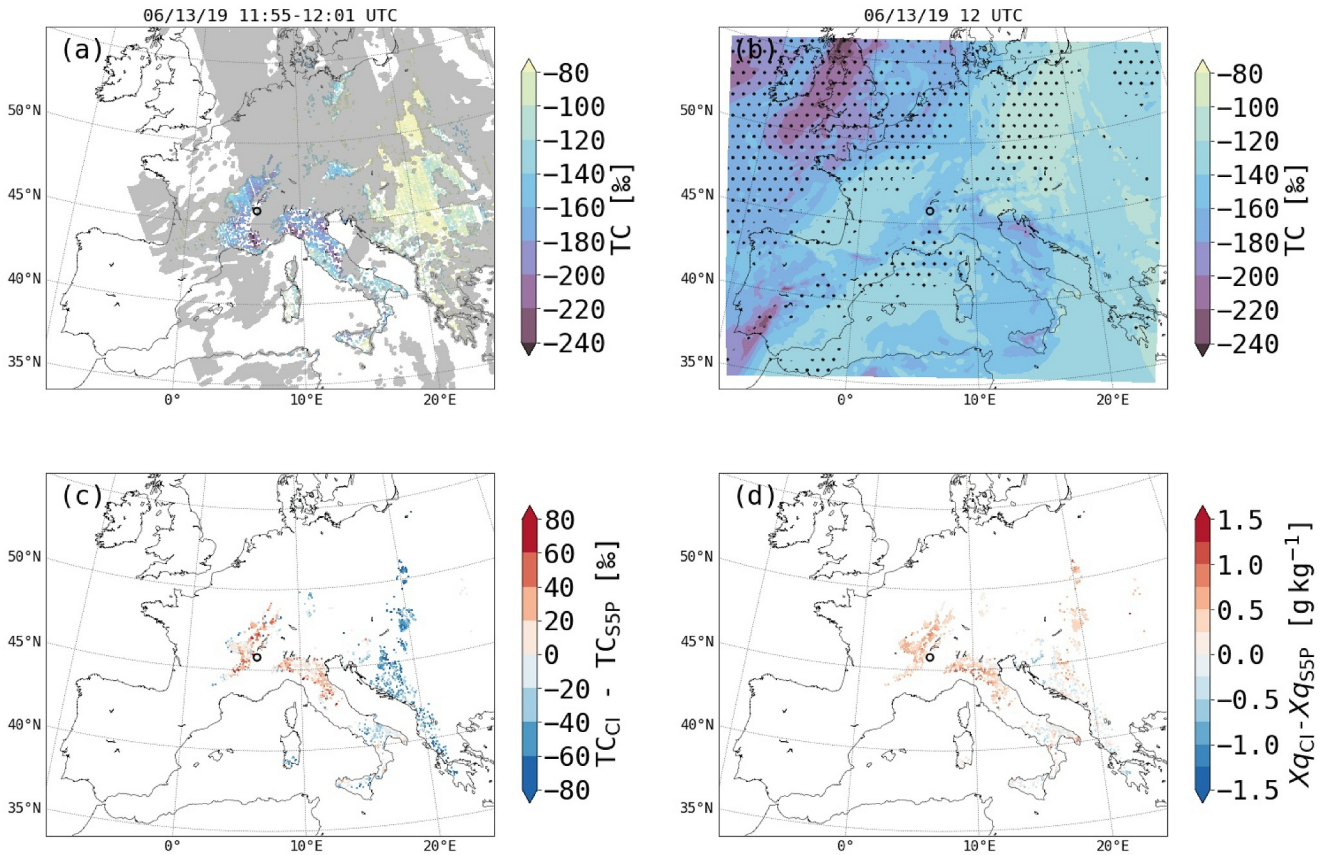

Figure 4. Maps of TC (a) from the satellite retrieval for the overpass and (b) in $COSMO_{iso}$ at 12 UTC 13/06/2019, and the difference in (c) TC and (d) total column averaged specific humidity Xq between the interpolated values to the satellite pixel center in $COSMO_{iso}$ and the satellite retrievals. Panels (c) and (d) show only pixels where there is a surface pressure difference between the TROPOMI retrieval and $COSMO_{iso}$ of less than 5 hPa, no cloud cover in $COSMO_{iso}$ (<1%) and a satellite viewing zenith angle <55°. The gray shaded regions in panel (a) show cloud covered areas from the S5P Level 2 NPP cloud product (Siddans & Smith, 2018). The dotted regions in panel (b) denote cloud coverage from the $COSMO_{iso}$ simulation. The location of the L-WAIVE campaign is marked with a black circle.

As a sensitivity analysis and to find an appropriate value for f, the effect of applying different collocation masks on the satellite retrieved TC is tested. Three simple collocation masks with a constant radius of 500, 200, and 50 km around Annecy, and the collocation masks for three values of f as shown in Figure 5 are applied. The masked TC distribution can vary strongly depending on the applied collocation masks (see Figure S6 in Supporting Information S1). The effect of the three different f masks on the mean δD of the selected satellite pixels is relatively small compared to the difference to the distribution for masks with a constant radius for most of the satellite overpasses. Based on this sensitivity analysis of masked mean δD, the fraction f used in this study at Annecy is set to 0.5 as this setup generates a collocation mask for each day representing well the daily δD variability. Compared to a constant collocation criteria of 50 km as used in the S5P-I validation report (M. Schneider et al., 2021), the number of collocated pixels increased on average by a factor of 20 when applying the collocation masks.

As illustrated in Figure 5, the size of the collocation mask changes with each overpass. This temporal variability can be explained by the dominant weather systems during the L-WAIVE campaign period. On 15 June 2019, due to a low pressure system over Germany, the collocation mask covers the largest region. The system leads to a consistent flow over a large area, which is reflected in high ρ over Southern France. The mean distance of points in the collocation mask from Annecy (Figure 6) before 16 June can exceed a distance of 500 km from Annecy. These large collocation areas reflect the aforementioned large-scale circulation pattern in the beginning of the campaign. After June 16, the local topography was an important factor of influence for local convective activity around Annecy, leading to small collocation areas.

10.1029/2023EA003400

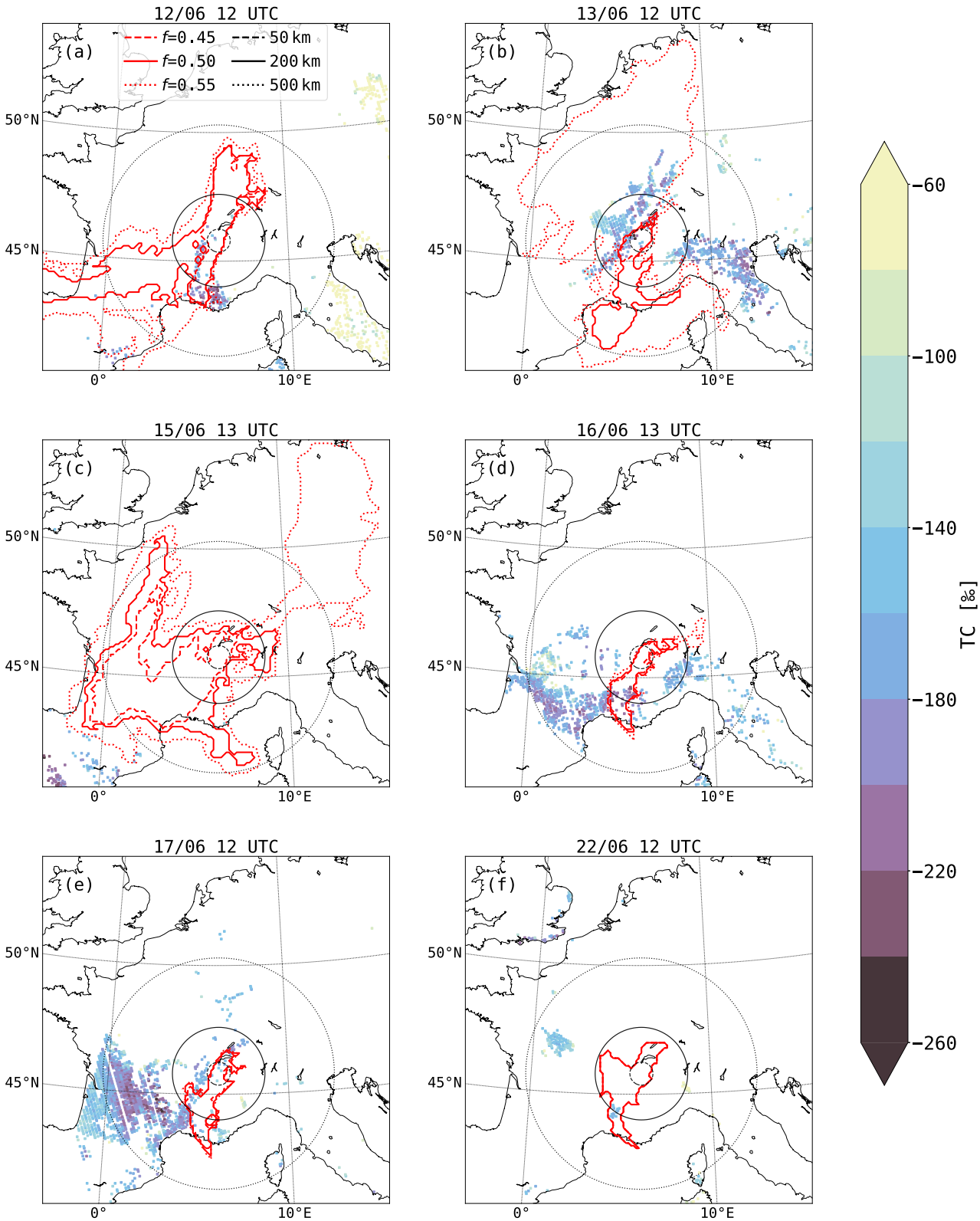

Figure 5. Satellite retrieved TC for six satellite overpasses during the L-WAIVE campaign. For each overpass, the collocation masks are shown for $f = 0.45$, 0.50, and 0.55 (red contours), and for a radius of 500 km (black dotted line), 200 km (black solid line) and 50 km (black dashed line). The satellite pixels are filtered to only include pixels that lie inside the $COSMO_{iso}$ domain, have a surface pressure difference between the TROPOMI retrieval and $COSMO_{iso}$ of less than 5 hPa, no cloud cover in $COSMO_{iso}$ (<1%) and a satellite viewing zenith angle <55°.

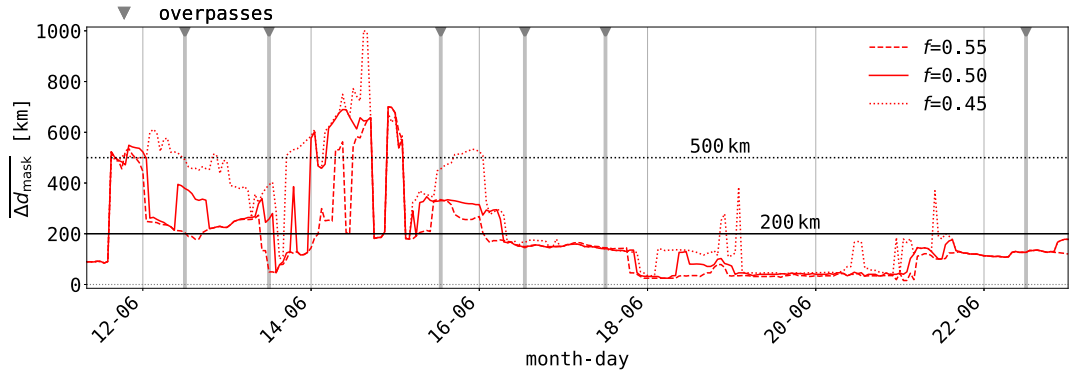

Figure 6. Time evolution of the median distance of the collocation region from Annecy for masks with $f = 0.45$, 0.50, and 0.55. A distance of 200 and 500 km is indicated by horizontal lines. The six satellite overpasses are marked by gray triangles with vertical lines.

5.2. Collocation Comparison Between In Situ and Satellite Data

In this section, we will apply the collocation regions to a 2 weeks case study in June 2019 to evaluate satellite retrieved TC with δD in atmospheric water vapor from the L-WAIVE in situ measurements and the $COSMO_{iso}$ simulation. There are two main questions to be answered: (a) How large is the observed difference between SC and TC? (b) What is the added value of TC retrievals compared to SC observations?

The overall comparison of SC and TC shows a similar temporal evolution of the two observational data sets (triangle markers in Figure 7a). There is an increase in TC and SC from 12 to 15 June, a local minimum around 16/17 June, followed by an increase until 20 June. To compare TC and SC, periods where satellite retrievals are available simultaneously with in situ measurements are of special interest. Satellite retrieved TC and L-WAIVE SC are available with less than 6 hr time shift on 13, 16, and 17 June. On 13 and 16 June, the difference between SC and TC is large with values up to 80‰. For 17 June, the difference between SC and TC is around 30‰. The $COSMO_{iso}$ vertical gradient for these days (Figure 7b) shows that high values of δD reach up to a height representing 95% of the total column water vapor on 17 June. On 13 and, especially, 16 June, there are low δD reaching below 700 hPa during the satellite overpasses. These 2 days were experiencing large-scale subsidence as also visible by the low cloud clover on these days. The simulated gradients indicate that the large SC—TC difference on 13 and 16 June originates from anomalously low δD in the middle and upper troposphere. The comparison of TC and SC observations can thus inform on vertical δD gradients, and the observed differences depend on the large-scale forcing.

The model output allows to investigate SC—TC differences in more detail. Based on the vertical δD distribution, column averaged δD has been calculated for different fractions of the total column water vapor content. Time series of SC representing 75%, 95%, and 100% (starting from the ground) of the total column water vapor content are shown in Figure 7a, labeled as SC_{75}, SC_{95}, and TC, respectively. SC_{75} corresponds approximately to the L-WAIVE upper flight height. Overall, the column averaged δD shows a similar temporal evolution for the three column averages marked by decreasing values with increasing column top height (Figure 7a). The difference, SC_{75} minus TC, derived from $COSMO_{iso}$ can be interpreted as the expected difference between the SC from the in situ measurements and the TC TROPOMI retrievals. The mean difference between SC_{75} and TC is 33.2‰ with a 25–75 percentile range of 23.5–40.1‰. The temporal variability of this difference reflects the changes between periods of large-scale subsidence and high-reaching boundary layer δD. For periods of large-scale subsidence under the influence of a high pressure system (e.g., in the afternoon of 16 June), SC_{75} increases while SC_{95} and TC decrease. For periods with high-reaching boundary layer δD (see Figure 7b), the temporal evolution of SC and TC agrees more closely (see e.g., 14 and 15 June). Thus, during large-scale subsidence, TCs can give additional information on the vertical state of the atmosphere. For example, the larger the difference between SC_{75} and TC during large-scale subsidence, the more depleted in deuterium is the upper troposphere. Thus, satellite retrieved TC can provide valuable additional information in combination with in situ measurements. Moreover, satellite retrievals are likely to be retrieved during large-scale subsidence due to low cloud coverage in such situations. At the same time, the increased difference between lower troposphere SC and TC in these situations makes the

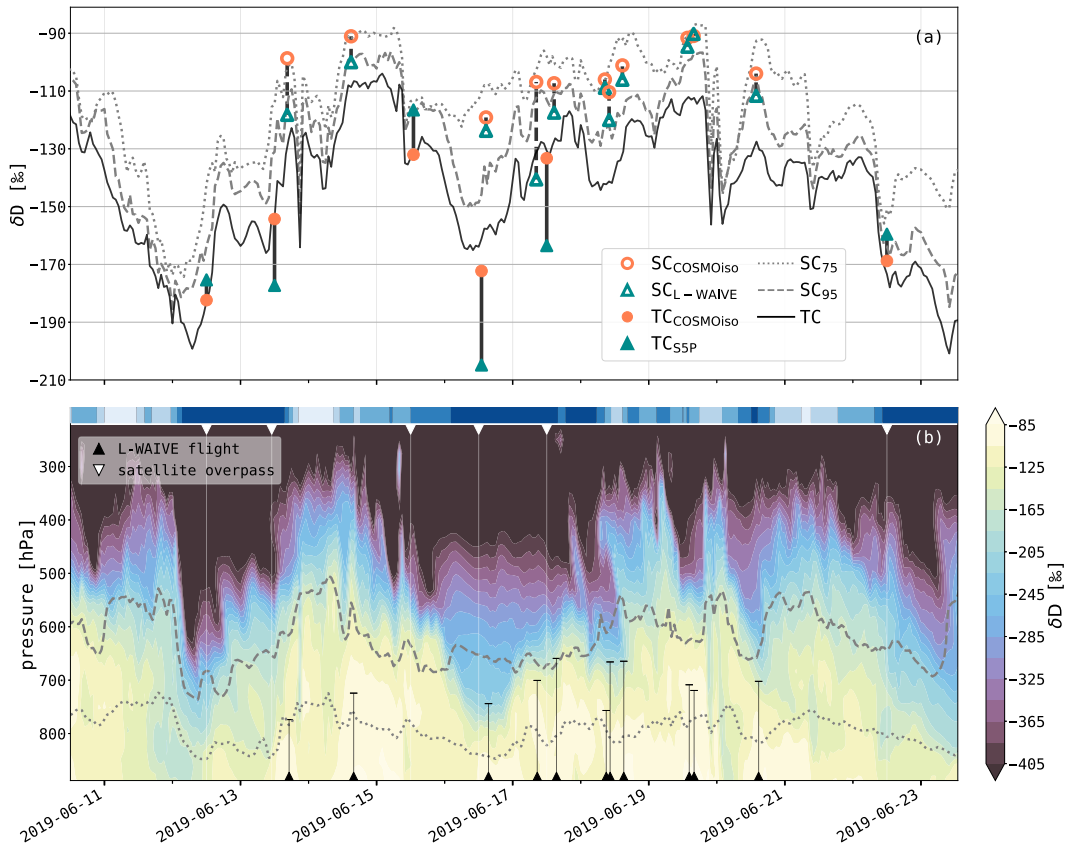

Figure 7. (a) Time evolution of SC from the L-WAIVE measurements (open green triangles) and $COSMO_{iso}$ profiles (open orange circles), and median TC in collocation mask $f = 0.5$ from satellite retrievals (filled green triangles) and $COSMO_{iso}$ (filled orange circles). The solid and dashed vertical lines connect corresponding measured and modeled TC and SC, respectively. The lines show TC (black solid), SC_{95} (gray dashed) and SC_{75} (gray dotted) at Annecy in $COSMO_{iso}$. (b) Vertical cross section of δD in $COSMO_{iso}$ at Annecy from 10 to 23 June 2019. The gray dotted and dashed lines show the height bounding 75% and 95%, respectively, of the vertically integrated water vapor content starting from the surface. The blue shading at the top marks the $COSMO_{iso}$ middle and high cloud coverage at Annecy (lightest blue representing 80%–100%, darkest blue 0%–20% cloud cover).

validation of satellite retrieved TC with in situ measurements more demanding as high altitude or complete column in situ measurements are needed for such purposes.

By comparing TC-SC in the observations and $COSMO_{iso}$, the origin of model biases can be identified. On 16 June and 17 June afternoon, while SC of $COSMO_{iso}$ and L-WAIVE correspond well, the difference between TC and SC is too small in $COSMO_{iso}$. The good agreement of the SC values indicates that $COSMO_{iso}$ underestimates the decrease in δD above 3,000 m during this period of large-scale subsidence. Therefore, the TC-SC difference does not only inform on the vertical δD gradient but can also relate model biases to specific vertical layers.

6. Discussion

6.1. From Subcolumns to Total Columns

The evaluation of satellite retrieved TC with in situ measured SC revealed a number of differences between the data sets. One of the main findings is that TC provides information on vertical δD gradients in combination with SC measurements. From the observational data sets of this study, we saw that the difference between TC and SC increases under large-scale subsidence when there is a steep δD gradient from the boundary layer into the free troposphere and an increased influence of upper tropospheric δD on TC. This illustrates that our collocation method enables a meaningful comparison of TC and SC observations and that simultaneous measurements of δD from various platforms at different levels and layers can help to characterize vertical δD profiles. TC

measurements can thus provide important information of vertical water vapor mixing into the free troposphere, especially in combination with SC measurements from in situ measurements or other satellite products (e.g., δD retrievals from the Infrared Atmospheric Sounding Interferometer on Metop-A and Metop-B with a maximum sensitivity at 4.2 km height; M. Schneider et al., 2016).

6.2. $COSMO_{iso}$ Simulation

Difference between modeled and observed δD in water vapor can help to identify model shortcomings. Overall, a positive δD bias was observed in $COSMO_{iso}$ for the collocation masks at Annecy. A possible explanation for too high boundary layer δD and q could be too weak vertical mixing in $COSMO_{iso}$ leading to too low upper level δD, while the boundary layer remains too enriched in heavy isotopes. The $COSMO_{iso}$ profiles follow a linear mixing behavior that is shifted relative to the observed profiles (see also Figure S5 in Supporting Information S1). This is also reflected by the large model bias across the boundary layer top above 800 hPa. While too weak mixing could partly explain the $COSMO_{iso}$ bias, a biased isotopic composition of the free troposphere can also lead to a shift in the mixing line in the q-δD-space. Further, the larger bias in $COSMO_{iso}$ for TC than for SC indicates that the positive δD bias is not only related to too high δD in the boundary layer but also to too high modeled δD in the upper troposphere. Due to limited observations of stable water isotopes in water vapor in the upper troposphere, isotope-enabled models are only weakly constrained at high altitudes leading to high uncertainties in modeled δD at these levels. The shown difference larger than 20‰ between modeled and observed TC during large-scale subsidence indicates the upper tropospheric δD is overestimated by more than 20‰ in $COSMO_{iso}$. The positive δD bias in the boundary layer might even lead to an underestimation of the upper-tropospheric negative δD bias.

7. Conclusions

In this study, we introduced a new collocation method to evaluate satellite retrieved total column averaged δD with in situ measurements different spatial and temporal scales. We translated in situ vertical δD profiles into subcolumn averages and selected satellite pixels within the in situ measurement's collocation regions. This allowed us to meaningfully compare in situ SC with satellite retrieved TC to evaluate the additional information gained from TC observations.

We show that we can use $COSMO_{iso}$ simulations to extrapolate and translate in situ δD measurements to use for the evaluation of the spatially and temporally coarser data set of satellite retrieved total column δD. The difference between SC and TC is largest during situations of large-scale subsidence and low boundary layer extent. As these situations often coincide with low cloud coverage, more satellite data are available during large-scale subsidence than during situation of frontal passages or local convective systems when the difference between SC and TC is smaller. Therefore, the combination of satellite retrieved TC with in situ boundary layer measurements allows for estimates of δD vertical gradients under large-scale subsidence. Such an improved understanding of vertical δD gradients is important as vertical δD gradients are a measure of the rain out fraction of an air parcel (as e.g., in the Rayleigh model by Dansgaard, 1964) and can be used to estimate vertical mixing (e.g., Benetti et al., 2018; Hu et al., 2022). Satellite retrieved TC provides thus a valuable global data set to report on δD vertical gradients and its variability.

The comparison of in situ measurements and TROPOMI retrieved TC with $COSMO_{iso}$ simulations reveals a positive model bias but a structural agreement between the data sets. The positive model bias has been attributed to too high modeled δD in (a) the planetary boundary layer and (b) the upper troposphere. The importance of the model bias at these two vertical layers for TC and SC depends on the large-scale meteorological situation with a stronger impact from upper tropospheric biases during large-scale subsidence. Our method provides a way for future studies to systematically investigate model bias in the upper tropospheric δD by combining in situ measurements and satellite retrievals.

These results lead to two main conclusions on the vertical scales: Satellite retrieved TC can give additional information on the vertical δD distribution in combination with boundary layer measurements of δD. And investigating the difference between SC and TC observations can be a useful tool to identify upper tropospheric δD biases in isotope-enabled numerical models.

This study cannot cover all potential aspects of TC evaluation as we are only showcasing our methodology with one case study of 2 weeks. For different locations and time periods, the results from such an evaluation can differ. The method introduced here should be applied to different locations (e.g., at different altitudes or climate zones) to investigate the variability of δD spatial representativeness, longer time series (e.g., using $COSMO_{iso}$ climate simulations) to study time scales beyond 2 weeks and other data sets (e.g., satellite retrievals during cloudy conditions or a combination of TROPOMI retrievals with the infrared atmospheric sounding interferometer (IASI) onboard the Metop-A and Metop-B platforms) to further improve our understanding of δD spatial scales.

In this study, we have demonstrated that our method provides a more robust statistical basis compared to a constant distance collocation criterion for the in situ evaluation of δD satellite retrievals. The method will thus be useful in planning and executing forthcoming validation campaigns by providing an estimate of the representativeness of the measurement location and the appropriate vertical extent of in situ measurements. Further, our method can potentially also be applied for the evaluation and validation of other satellite products.

Appendix A: Definition of the Pressure Weighting Function for Calculating Column Averaged H_2O

The calculation of the column averaged δD follows O'Dell et al. (2012) using a pressure weighting function h'. To calculate h', the atmosphere is divided into z pressure levels that define $z-1$ pressure layers. In the following, if v_i denotes a variable on pressure level i, v_j is the average value of pressure layer j. The column averaged water vapor mole mixing ratio XH_2O is calculated by applying a pressure weighting function h' to the mean layer water vapor mixing ratio m'_{H2O}:

$$XH_2O = \sum_{i=1}^{z-1} h'_i m'_{H2O,i}, \tag{A1}$$

with

$$m'_{H2O,i} = (1 - f_i) \cdot m_{H2O,i} + f_i \cdot m_{H2O,i+1}, \tag{A2}$$

and the pressure weighting function

$$h'_i = \frac{c_i' \Delta p_i}{\sum_{i=1}^{z-1} c_i' \Delta p_i}, \tag{A3}$$

where $c'_i = 0.5 \cdot c_i + 0.5 \cdot c_{i+1}$ is the density of the air per unit pressure of pressure layer i with $c_i = \frac{1-q_i}{gM_{dry}}$. q_i is the specific humidity at pressure level i, g the local gravity acceleration, M_{dry} the molar mass of dry air and Δp_i is the pressure difference between pressure level i and $i+1$. f_i is a weighting factor for each pressure level and equal to 0.5, except for the lowest pressure level 0 where $f_s = \frac{p_S - p_0}{p_1 - p_0}$. This formulation assumes that changes in each bin are linear with height and that the surface pressure (p_S) is contained in the lowermost layer confined by pressure levels 0 and 1.

Data Availability Statement

The data presented in this paper can be obtained as follows:

- The L-WAIVE vertical profiles are published on Zenodo with the following doi: https://doi.org/10.5281/zenodo.8430279 (Sodemann & Seidl, 2023).
- The presented $COSMO_{iso}$ interpolated profiles, subcolumn averaged δD and the model output compared to the satellite overpasses in Figure 4, Figures S1 and S2 in Supporting Information S1 are published on Zenodo (https://doi.org/10.5281/zenodo.8430434; Thurnherr & Sodemann, 2023). The full $COSMO_{iso}$ simulation has a too large size to be published on a public online repository and will be provided by the authors upon request.

- The TROPOMI L2 Prototype Stable Water Vapor Isotopologue Product (V1.0.0) for June 2019 is available on Zenodo with the following doi: https://doi.org/10.5281/zenodo.10877376 (Trent & Boesch, 2024).

Acknowledgments
This study was funded by the Sentinel-5P + Innovation Water Vapour Isotopologues (H2O-ISO) project (ESA/Contract No.4000127561/19/I-NS). The $COSMO_{iso}$ simulations were performed on the Norwegian high-performance computing infrastructure SIGMA2. The collocation mask was coded in python and inspired by the ConTrack python modul (https://github.com/steidani/ConTrack). We thank Matthias Schneider, Farahnaz Koshrawi, Amélie N. Röhling and Christopher Diekmann for the inspiring exchange during the H2O-ISO project meetings. We thankfully acknowledge the two anonymous reviewers and the editor Cinzia Cervato for their detailed and constructive comments that helped to improve the paper.

References

Aemisegger, F., Spiegel, J. K., Pfahl, S., Sodemann, H., Eugster, W., & Wernli, H. (2015). Isotope meteorology of cold front passages: A case study combining observations and modeling: Water isotopes during cold fronts. *Geophysical Research Letters*, *42*(13), 5652–5660. https://doi.org/10.1002/2015GL063988

Aemisegger, F., Sturm, P., Graf, P., Sodemann, H., Pfahl, S., Knohl, A., & Wernli, H. (2012). Measuring variations of $\delta^{18}O$ and $\delta^{2}H$ in atmospheric water vapour using two commercial laser-based spectrometers: An instrument characterisation study. *Atmospheric Measurement Techniques*, *5*(7), 1491–1511. https://doi.org/10.5194/amt-5-1491-2012

Bailey, A., Noone, D., Cobb, K., Atwood, A., Dee, S., & Nusbaumer, J. (2021). Water isotopes and climate US CLIVAR Report, 2021-1. https://doi.org/10.5065/rmyf-qw78

Benetti, M., Lacour, J.-L., Sveinbjörnsdóttir, Á. E., Aloisi, G., Reverdin, G., Risi, C., et al. (2018). A framework to study mixing processes in the marine boundary layer using water vapor isotope measurements. *Geophysical Research Letters*, *45*(5), 2524–2532. https://doi.org/10.1002/2018GL077167

Benetti, M., Steen-Larsen, H. C., Reverdin, G., Sveinbjörnsdóttir, Á. E., Aloisi, G., Berkelhammer, M. B., et al. (2017). Stable isotopes in the atmospheric marine boundary layer water vapour over the Atlantic Ocean, 2012–2015. *Scientific Data*, *4*(1), 160128. https://doi.org/10.1038/sdata.2016.128

Blossey, P. N., Kuang, Z., & Romps, D. M. (2010). Isotopic composition of water in the tropical tropopause layer in cloud-resolving simulations of an idealized tropical circulation. *Journal of Geophysical Research*, *115*(D24). https://doi.org/10.1029/2010JD014554

Boesch, H., Deutscher, N. M., Warneke, T., Byckling, K., Cogan, A. J., Griffith, D. W. T., et al. (2013). HDO/H_2O ratio retrievals from GOSAT. *Atmospheric Measurement Techniques*, *6*(3), 599–612. https://doi.org/10.5194/amt-6-599-2013

Bonne, J.-L., Behrens, M., Meyer, H., Kipfstuhl, S., Rabe, B., Schönicke, L., et al. (2019). Resolving the controls of water vapour isotopes in the Atlantic sector. *Nature Communications*, *10*(1), 1632. https://doi.org/10.1038/s41467-019-09242-6

Brady, E., Stevenson, S., Bailey, D., Liu, Z., Noone, D., Nusbaumer, J., et al. (2019). The connected isotopic water cycle in the community Earth system model version 1. *Journal of Advances in Modeling Earth Systems*, *11*(8), 2547–2566. https://doi.org/10.1029/2019MS001663

Cauquoin, A., & Werner, M. (2021). High-resolution nudged isotope modeling with ECHAM6-wiso: Impacts of updated model physics and ERA5 reanalysis data. *Journal of Advances in Modeling Earth Systems*, *13*(11), e2021MS002532. https://doi.org/10.1029/2021MS002532

Chazette, P., Flamant, C., Sodemann, H., Totems, J., Monod, A., Dieudonné, E., et al. (2021). Experimental investigation of the stable water isotope distribution in an Alpine lake environment (L-WAIVE). *Atmospheric Chemistry and Physics*, *21*(14), 10911–10937. https://doi.org/10.5194/acp-21-10911-2021

Cogan, A. J., Boesch, H., Parker, R. J., Feng, L., Palmer, P. I., Blavier, J.-F. L., et al. (2012). Atmospheric carbon dioxide retrieved from the Greenhouse gases Observing SATellite (GOSAT): Comparison with ground-based TCCON observations and GEOS-Chem model calculations. *Journal of Geophysical Research*, *117*(D21). https://doi.org/10.1029/2012JD018087

Craig, H., & Gordon, L. (1965). Deuterium and oxygen 18 variations in the ocean and the marine atmosphere. In *Proceedings of the Stable Isotopes in Oceanographic Studies and Paleotemperatures*.

Dahinden, F., Aemisegger, F., Wernli, H., Schneider, M., Diekmann, C. J., Ertl, B., et al. (2021). Disentangling different moisture transport pathways over the eastern subtropical north Atlantic using multi-platform isotope observations and high-resolution numerical modelling. *Atmospheric Chemistry and Physics*, *21*(21), 16319–16347. https://doi.org/10.5194/acp-21-16319-2021

Dansgaard, W. (1964). Stable isotopes in precipitation. *Tellus*, *16*(4), 436–468. https://doi.org/10.3402/tellusa.v16i4.8993

de Vries, A. J., Aemisegger, F., Pfahl, S., & Wernli, H. (2022). Stable water isotope signals in tropical ice clouds in the west African monsoon simulated with a regional convection-permitting model. *Atmospheric Chemistry and Physics*, *22*(13), 8863–8895. https://doi.org/10.5194/acp-22-8863-2022

Diekmann, C. J., Schneider, M., Knippertz, P., de Vries, A. J., Pfahl, S., Aemisegger, F., et al. (2021). A Lagrangian perspective on stable water isotopes during the west African monsoon. *Journal of Geophysical Research: Atmospheres*, *126*(19), e2021JD034895. https://doi.org/10.1029/2021JD034895

Dütsch, M., Pfahl, S., Meyer, M., & Wernli, H. (2018). Lagrangian process attribution of isotopic variations in near-surface water vapour in a 30-year regional climate simulation over Europe. *Atmospheric Chemistry and Physics*, *18*(3), 1653–1669. https://doi.org/10.5194/acp-18-1653-2018

Frankenberg, C., Wunch, D., Toon, G., Risi, C., Scheepmaker, R., Lee, J.-E., et al. (2013). Water vapor isotopologue retrievals from high-resolution GOSAT shortwave infrared spectra. *Atmospheric Measurement Techniques*, *6*(2), 263–274. https://doi.org/10.5194/amt-6-263-2013

Frankenberg, C., Yoshimura, K., Warneke, T., Aben, I., Butz, A., Deutscher, N., et al. (2009). Dynamic processes governing lower-tropospheric HDO/H_2O ratios as observed from space and ground. *Science*, *325*(5946), 1374–1377. https://doi.org/10.1126/science.1173791

Galewsky, J., Steen-Larsen, H. C., Field, R. D., Worden, J., Risi, C., & Schneider, M. (2016). Stable isotopes in atmospheric water vapor and applications to the hydrologic cycle. *Reviews of Geophysics*, *54*(4), 809–865. https://doi.org/10.1002/2015RG000512

Gat, J. R. (1996). Oxygen and hydrogen isotopes in the hydrologic cycle. *Annual Review of Earth and Planetary Sciences*, *24*(1), 225–262. https://doi.org/10.1146/annurev.earth.24.1.225

Gimeno, L., Eiras-Barca, J., Durán-Quesada, A. M., Dominguez, F., Van Der Ent, R., Sodemann, H., et al. (2021). The residence time of water vapour in the atmosphere. *Nature Reviews Earth & Environment*, *2*(8), 558–569. https://doi.org/10.1038/s43017-021-00181-9

Guerlet, S., Butz, A., Schepers, D., Basu, S., Hasekamp, O. P., Kuze, A., et al. (2013). Impact of aerosol and thin cirrus on retrieving and validating XCO2 from GOSAT shortwave infrared measurements. *Journal of Geophysical Research: Atmospheres*, *118*(10), 4887–4905. https://doi.org/10.1002/jgrd.50332

Henze, D., Noone, D., & Toohey, D. (2022). Aircraft measurements of water vapor heavy isotope ratios in the marine boundary layer and lower troposphere during oracles. *Earth System Science Data*, *14*(4), 1811–1829. https://doi.org/10.5194/essd-14-1811-2022

Herbin, H., Hurtmans, D., Turquety, S., Wespes, C., Barret, B., Hadji-Lazaro, J., et al. (2007). Global distributions of water vapour isotopologues retrieved from IMG/ADEOS data. *Atmospheric Chemistry and Physics*, *7*(14), 3957–3968. https://doi.org/10.5194/acp-7-3957-2007

Herman, R. L., Cherry, J. E., Young, J., Welker, J. M., Noone, D., Kulawik, S. S., & Worden, J. (2014). Aircraft validation of aura tropospheric emission spectrometer retrievals of HDO/H_2O. *Atmospheric Measurement Techniques*, *7*(9), 3127–3138. https://doi.org/10.5194/amt-7-3127-2014

Herman, R. L., Worden, J., Noone, D., Henze, D., Bowman, K., Cady-Pereira, K., et al. (2020). Comparison of optimal estimation HDO/H_2O retrievals from AIRS with ORACLES measurements. *Atmospheric Measurement Techniques*, *13*(4), 1825–1834. https://doi.org/10.5194/amt-13-1825-2020

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Hu, J., Bailey, A., Nusbaumer, J., Dee, S., Sasser, C., & Worden, J. (2022). Tracking shallow convective mixing and its influence on low-level clouds with stable water isotopes in vapor. *Journal of Geophysical Research: Atmospheres*, *127*(5), e2021JD035355. https://doi.org/10.1029/2021JD035355

KNMI. (2017). Algorithm theoretical basis document for the TROPOMI L01b data processor Tech. Rep. S5P-KNMI-L01B-0009-SD, CI-6480-ATBD, issue 8.0.0. Retrieved from https://sentinels.copernicus.eu/documents/247904/2476257/Sentinel-5P-TROPOMI-Level-1B-ATBD

Lacour, J.-L., Risi, C., Clarisse, L., Bony, S., Hurtmans, D., Clerbaux, C., & Coheur, P.-F. (2012). Mid-tropospheric δD observations from IASI/MetOp at high spatial and temporal resolution. *Atmospheric Chemistry and Physics*, *12*(22), 10817–10832. https://doi.org/10.5194/acp-12-10817-2012

Lacour, J.-L., Risi, C., Worden, J., Clerbaux, C., & Coheur, P.-F. (2018). Importance of depth and intensity of convection on the isotopic composition of water vapor as seen from IASI and TES δD observations. *Earth and Planetary Science Letters*, *481*, 387–394. https://doi.org/10.1016/j.epsl.2017.10.048

Landwehr, S., Volpi, M., Haumann, F. A., Robinson, C. M., Thurnherr, I., Ferracci, V., et al. (2021). Exploring the coupled ocean and atmosphere system with a data science approach applied to observations from the Antarctic Circumnavigation Expedition. *Earth System Dynamics*, *12*(4), 1295–1369. https://doi.org/10.5194/esd-12-1295-2021

Lee, J., Worden, J., Noone, D., Bowman, K., Eldering, A., LeGrande, A., et al. (2011). Relating tropical ocean clouds to moist processes using water vapor isotope measurements. *Atmospheric Chemistry and Physics*, *11*(2), 741–752. https://doi.org/10.5194/acp-11-741-2011

Lee, K.-O., Aemisegger, F., Pfahl, S., Flamant, C., Lacour, J.-L., & Chaboureau, J.-P. (2019). Contrasting stable water isotope signals from convective and large-scale precipitation phases of a heavy precipitation event in southern Italy during HyMeX IOP 13: A modelling perspective. *Atmospheric Chemistry and Physics*, *19*(11), 7487–7506. https://doi.org/10.5194/acp-19-7487-2019

Noone, D. (2012). Pairing measurements of the water vapor isotope ratio with humidity to deduce atmospheric moistening and dehydration in the tropical midtroposphere. *Journal of Climate*, *25*(13), 4476–4494. https://doi.org/10.1175/JCLI-D-11-00582.1

O'Dell, C. W., Connor, B., Bösch, H., O'Brien, D., Frankenberg, C., Castano, R., et al. (2012). The ACOS CO_2 retrieval algorithm—Part 1: Description and validation against synthetic observations. *Atmospheric Measurement Techniques*, *5*(1), 99–121. https://doi.org/10.5194/amt-5-99-2012

Oshchepkov, S., Bril, A., Yokota, T., Morino, I., Yoshida, Y., Matsunaga, T., et al. (2012). Effects of atmospheric light scattering on spectroscopic observations of greenhouse gases from space: Validation of PPDF-based CO_2 retrievals from GOSAT. *Journal of Geophysical Research*, *117*(D12). https://doi.org/10.1029/2012JD017505

Pfahl, S., & Wernli, H. (2008). Air parcel trajectory analysis of stable isotopes in water vapor in the eastern Mediterranean. *Journal of Geophysical Research*, *113*(D20), D20104. https://doi.org/10.1029/2008JD009839

Pfahl, S., Wernli, H., & Yoshimura, K. (2012). The isotopic composition of precipitation from a winter storm—A case study with the limited-area model COSMOiso. *Atmospheric Chemistry and Physics*, *12*(3), 1629–1648. https://doi.org/10.5194/acp-12-1629-2012

Risi, C., Bony, S., Vimeux, F., & Jouzel, J. (2010). Water-stable isotopes in the LMDZ4 general circulation model: Model evaluation for present-day and past climates and applications to climatic interpretations of tropical isotopic records. *Journal of Geophysical Research*, *115*(D12). https://doi.org/10.1029/2009JD013255

Risi, C., Noone, D., Worden, J., Frankenberg, C., Stiller, G., Kiefer, M., et al. (2012). Process-evaluation of tropospheric humidity simulated by general circulation models using water vapor isotopologues: 1. Comparison between models and observations. *Journal of Geophysical Research*, *117*(D5). https://doi.org/10.1029/2011JD016621

Rokotyan, N. V., Zakharov, V. I., Gribanov, K. G., Schneider, M., Bréon, F.-M., Jouzel, J., et al. (2014). A posteriori calculation of $\delta^{18}O$ and δD in atmospheric water vapour from ground-based near-infrared FTIR retrievals of $H_2{}^{16}O$, $H_2{}^{18}O$, and $HD^{16}O$. *Atmospheric Measurement Techniques*, *7*(8), 2567–2580. https://doi.org/10.5194/amt-7-2567-2014

Salmon, O. E., Welp, L. R., Baldwin, M. E., Hajny, K. D., Stirm, B. H., & Shepson, P. B. (2019). Vertical profile observations of water vapor deuterium excess in the lower troposphere. *Atmospheric Chemistry and Physics*, *19*(17), 11525–11543. https://doi.org/10.5194/acp-19-11525-2019

Schneider, A., Borsdorff, T., aan de Brugh, J., Aemisegger, F., Feist, D. G., Kivi, R., et al. (2020). First data set of H_2O/HDO columns from the tropospheric monitoring instrument (TROPOMI). *Atmospheric Measurement Techniques*, *13*(1), 85–100. https://doi.org/10.5194/amt-13-85-2020

Schneider, M., Barthlott, S., Hase, F., González, Y., Yoshimura, K., García, O. E., et al. (2012). Ground-based remote sensing of tropospheric water vapour isotopologues within the project MUSICA. *Atmospheric Measurement Techniques*, *5*(12), 3007–3027. https://doi.org/10.5194/amt-5-3007-2012

Schneider, M., González, Y., Dyroff, C., Christner, E., Wiegele, A., Barthlott, S., et al. (2015). Empirical validation and proof of added value of musica's tropospheric δD remote sensing products. *Atmospheric Measurement Techniques*, *8*(1), 483–503. https://doi.org/10.5194/amt-8-483-2015

Schneider, M., Hase, F., & Blumenstock, T. (2006). Ground-based remote sensing of HDO/H_2O ratio profiles: Introduction and validation of an innovative retrieval approach. *Atmospheric Chemistry and Physics*, *6*(12), 4705–4722. https://doi.org/10.5194/acp-6-4705-2006

Schneider, M., Röhling, A., Diekmann, C., Trent, T., Boesch, H., & Sodemann, H. (2021). *Sentinel-5p+ innovation (S5p+I) -water vapour isotopologues (H2O-ISO) validation report (VR)*. European Space Agency. Retrieved from https://s5pinnovationh2o-iso.le.ac.uk/wp-content/uploads/2021/10/S5p-I_VR_Version1.3.pdf

Schneider, M., Wiegele, A., Barthlott, S., González, Y., Christner, E., Dyroff, C., et al. (2016). Accomplishments of the MUSICA project to provide accurate, long-term, global and high-resolution observations of tropospheric H_2O, δD pairs—A review. *Atmospheric Measurement Techniques*, *9*(7), 2845–2875. https://doi.org/10.5194/amt-9-2845-2016

Siddans, R., & Smith, A. (2018). *Sentinel-5 precursor/TROPOMI level 2 product user manual NPP cloud*. European Space Agency. Retrieved from https://sentinels.copernicus.eu/documents/247904/2474726/Sentinel-5P-Level-2-Product-User-Manual-NPP-Cloud-product

Sodemann, H. (2020). Beyond turnover time: Constraining the lifetime distribution of water vapor from simple and complex approaches. *Journal of the Atmospheric Sciences*, *77*(2), 413–433. https://doi.org/10.1175/JAS-D-18-0336.1

Sodemann, H., Aemisegger, F., Pfahl, S., Bitter, M., Corsmeier, U., Feuerle, T., et al. (2017). The stable isotopic composition of water vapour above Corsica during the HyMeX SOP1 campaign: Insight into vertical mixing processes from lower-tropospheric survey flights. *Atmospheric Chemistry and Physics*, *17*(9), 6125–6151. https://doi.org/10.5194/acp-17-6125-2017

Sodemann, H., & Seidl, A. (2023). Vertical profiles of stable water isotopes and thermodynamic properties from research flights during the L-WAIVE field campaign in June 2019. *Zenodo*. https://doi.org/10.5281/zenodo.8430279

Steen-Larsen, H. C., Sveinbjörnsdottir, Á. E., Peters, A. J., Masson-Delmotte, V., Guishard, M. P., Hsiao, G., et al. (2014). Climatic controls on water vapor deuterium excess in the marine boundary layer of the North Atlantic based on 500 days of in situ, continuous measurements. *Atmospheric Chemistry and Physics*, *14*(15), 7741–7756. https://doi.org/10.5194/acp-14-7741-2014

Steppeler, J., Doms, G., Schättler, U., Bitzer, H. W., Gassmann, A., Damrath, U., & Gregoric, G. (2003). Meso-gamma scale forecasts using the nonhydrostatic model LM. *Meteorology and Atmospheric Physics*, *82*(1), 75–96. https://doi.org/10.1007/s00703-001-0592-9

Thurnherr, I., Hartmuth, K., Jansing, L., Gehring, J., Boettcher, M., Gorodetskaya, I., et al. (2021). 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 and Climate Dynamics*, *2*(2), 331–357. https://doi.org/10.5194/wcd-2-331-2021

Thurnherr, I., Kozachek, A., Graf, P., Weng, Y., Bolshiyanov, D., Landwehr, S., et al. (2020). Meridional and vertical variations of the water vapour isotopic composition in the marine boundary layer over the Atlantic and Southern Ocean. *Atmospheric Chemistry and Physics*, *20*(9), 5811–5835. https://doi.org/10.5194/acp-20-5811-2020

Thurnherr, I., & Sodemann, H. (2023). Numerical weather simulation using COSMOiso in June 2019 during L-WAIVE field campaign: Selected model output and post-processed data. *Zenodo*. https://doi.org/10.5281/zenodo.8430434

Trent, T., & Boesch, H. (2024). University of leicester TROPOMI stable water vapour isotopologue (H2O-ISO) prototype product (Vesrion 1.0.0). *Zenodo*. https://doi.org/10.5281/zenodo.10877376

Trent, T., Boesch, H., Schneider, M., Röhling, A., Khosrawi, F., Diekmann, C., et al. (2021). *Sentinel-5p+ innovation (S5p+I) -water vapour isotopologues (H2O-ISO) algorithm theoretical basis document (ATBD)*. European Space Agency. Retrieved from https://s5pinnovationh2o-iso.le.ac.uk/wp-content/uploads/2021/10/S5P-I_ISO_ATBD_Version1.4.pdf

Trent, T., Boesch, H., Somkuti, P., & Scott, N. A. (2018). Observing water vapour in the planetary boundary layer from the short-wave infrared. *Remote Sensing*, *10*(9), 1469. https://doi.org/10.3390/rs10091469

Vergara-Temprado, J., Ban, N., Panosetti, D., Schlemmer, L., & Schär, C. (2019). Climate models permit convection at much coarser resolutions than previously considered. *Journal of Climate*, *33*(5), 1915–1933. https://doi.org/10.1175/JCLI-D-19-0286.1

Weng, Y., Johannessen, A., & Sodemann, H. (2021). High-resolution stable isotope signature of a land-falling atmospheric river in southern Norway. *Weather and Climate Dynamics*, *2*(3), 713–737. https://doi.org/10.5194/wcd-2-713-2021

Worden, J., Noone, D., & Bowman, K. (2007). Importance of rain evaporation and continental convection in the tropical water cycle. *Nature*, *445*(7127), 528–532. https://doi.org/10.1038/nature05508

Wunch, D., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Stephens, B. B., Fischer, M. L., et al. (2010). Calibration of the total carbon column observing network using aircraft profile data. *Atmospheric Measurement Techniques*, *3*(5), 1351–1362. https://doi.org/10.5194/amt-3-1351-2010

Wunch, D., Wennberg, P. O., Toon, G. C., Connor, B. J., Fisher, B., Osterman, G. B., et al. (2011). A method for evaluating bias in global measurements of CO_2 total columns from space. *Atmospheric Chemistry and Physics*, *11*(23), 12317–12337. https://doi.org/10.5194/acp-11-12317-2011

Yoshimura, K., Kanamitsu, M., Noone, D., & Oki, T. (2008). Historical isotope simulation using reanalysis atmospheric data. *Journal of Geophysical Research*, *113*(D19). https://doi.org/10.1029/2008JD010074