Assessing the potential of free tropospheric water vapour isotopologue satellite observations for improving the analyses of latent heating events

Matthias Schneider[1], Kinya Toride[2,3,a,b], Farahnaz Khosrawi[1,c], Frank Hase[1], Benjamin Ertl[1,4], Christopher J. Diekmann[1,d], and Kei Yoshimura[2,5]

[1]Institute of Meteorology and Climate Research (IMK-ASF), Karlsruhe Institute of Technology, Karlsruhe, Germany
[2]Institute for Industrial Science, The University of Tokyo, Chiba, Japan
[3]Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA
[4]Steinbuch Centre for Computing (SCC), Karlsruhe Institute of Technology, Karlsruhe, Germany
[5]Earth Observation Research Center, Japan Aerospace Exploration Agency, Japan
[a]now at: Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, Colorado, USA
[b]now at: NOAA Physical Sciences Laboratory, Boulder, Colorado, USA
[c]now at: Forschungszentrum Jülich GmbH, Jülich, Germany
[d]now at: Telespazio Germany GmbH, Darmstadt, Germany

Correspondence: M. Schneider
(matthias.schneider@kit.edu)

Abstract. Satellite-based observations of free tropospheric water vapour isotopologue ratios (δD) with good global and temporal coverage have become recently available. We investigate the potential of these observations for constraining the uncertainties of the atmospheric analyses fields of specific humidity (q), temperature (T), and δD and of variables that capture important properties of the atmospheric water cycle, namely the vertical velocity (ω), the latent heating rate (Q_2), and the precipitation rate (Prcp). Our focus is on the impact of the δD observations if used in addition to the observation of q and T, which are much easier to be observed by satellites and routinely in use for atmospheric analyses. For our investigations we use an Observing System Simulation Experiment, i.e. simulate the satellite observations of q, T, and δD with known uncertainties, then use them within a Kalman filter based assimilation framework in order to evaluate their potential for improving the quality of atmospheric analyses. The study is made for low latitudes ($30°$S to $30°$N) and for 40 days between mid-July and end of August 2016. We find that the assimilation of q and T observations alone well constrains the atmospheric q and T fields (analyses skills in the free troposphere of up to 60%), and moderately constrains the fields of δD, ω, Q_2, and Prcp (analyses skills of 20%-40%). The additional assimilation of δD observations further improves the quality of the analyses of all variables. We use Q_2 as proxy for the presence of condensation and evaporation processes, and we show that the additional improvement is rather weak when evaporation or condensation are negligible (additional analyses skills of generally below 5%), and strongest for high condensation rates (additional skills of about 15% and above). The very high condensation rates (identified by large positive Q_2 values) are rare, but related to extreme events (very high ω and Prcp) that are not well captured in the analyses (for these extreme events also the analyses uncertainties of ω, Q_2, and Prcp are very large), i.e. the additional assimilation of

δD observations significantly improves the analyses of the water cycle related variables for the events when an improvement is most important. In real world satellite datasets δD observations affected by such strong latent heating events are frequently

20 available, suggesting that the here demonstrated additional δD impact for the simulated world is also a realistic scenario for a real world data assimilation.

1 Introduction

Clouds and water vapour control atmospheric radiative heating/cooling and condensation or evaporation of water determines where latent heating or latent heat consumption takes place. The heating patterns then drive atmospheric circulation, thereby

25 causing additional evaporation/condensation and impacting on the distribution of water vapour and clouds, which in turn again modify the latent and radiative heating patterns of the atmosphere. This strong coupling between moisture pathways, diabatic heating and atmospheric circulation is responsible for important climate feedback mechanisms (e.g. Sherwood et al., 2014; Bony et al., 2015) and often is connected to the evolution of severe weather events (e.g. Fink et al., 2012; Evans et al., 2017). In this context, it is rather worrisome that the diabatic heating rates obtained from different current global reanalyses show

30 significant inconsistencies (e.g. Chan and Nigam, 2009; Ling and Zhang, 2013).

For the generation of daily and global scale analyses, the operational assimilaton systems assimilate the outgoing microwave or infrared radiation (e.g. Eyre et al., 2022). This radiation contains information on the atmospheric state (mostly atmospheric specific humidity, q, and temperature, T). There are many different satellites that measure this spectrally-resolved radiation, including operational weather satellites, like the European Meteosat- and Metop-series (https://www.eumetsat.int/our-satellites/

35 meteosat-series and https://www.eumetsat.int/our-satellites/metop-series, respectively).

In this study we investigate the information that free tropospheric δD isotopologue ratio observations can provide in addition to the information provided by the observations of q and T for improving the analyses. The ratio is calculated between the heavy isotopologue HD^{16}O and the main isotopologue H$_2^{16}$O in the vapour phase

$$\delta\text{D} = \frac{\text{HD}^{16}\text{O}/\text{H}_2^{16}\text{O}}{R_{\text{VSMOW}}} - 1, \qquad (1)$$

40 where H$_2^{16}$O and HD^{16}O are the concentrations of the respective isotopologues and $R_{\text{VSMOW}} = 3.1152 \times 10^{-4}$ is the Vienna Standard Mean Ocean Water ratio of the two isotopologues (this is a standard ratio typically encountered in Ocean water, Craig, 1961). Our particular interest is in δD, because it is very promising for improving the analysed latent heating rates: firstly, it can be observed with a reasonable precision by satellite on a daily and almost global scale (e.g. Diekmann et al., 2021a) and secondly, it is affected by condensation/evaporation and thus closely linked to latent heating processes (e.g. Galewsky et al.,

45 2016).

During the last 15 years, tropospheric δD products have been developed for different satellite sensors (e.g. Worden et al., 2007; Frankenberg et al., 2009; Schneider and Hase, 2011; Lacour et al., 2012; Boesch et al., 2013; Worden et al., 2019; Schneider et al., 2020). Meanwhile, different weather and climate models have the water isotopologues and the relevant physical processes implemented and can provide modelled isotopologue fields on a global and regional scale at different horizontal

50 resolutions (e.g. Yoshimura et al., 2008; Risi et al., 2010; Werner et al., 2011; Pfahl et al., 2012; Eckstein et al., 2018). This offers advanced opportunities for studying atmospheric moisture processes with water isotopologues.

The tropospheric water vapour isotopologue composition has been used for investigating water cycle related biases in atmospheric models (e.g. Risi et al., 2012; Field et al., 2014; Schneider et al., 2017), clouds and/or precipitation involving processes (e.g. Webster and Heymsfield, 2003; Worden et al., 2007; Blossey et al., 2010; Field et al., 2010; Bailey et al., 2015; Diek-

55 mann et al., 2021b), local diurnal-scale moisture transport (Noone et al., 2011; González et al., 2016), and large-scale moisture transport (e.g. Noone, 2012; González et al., 2016; Lacour et al., 2017; Dahinden et al., 2021).

We use a data assimilation framework together with an OSSE (Observation System Simulation Experiment) to document the added value of the free tropospheric δD satellite observations, i.e. we simulate satellite observations and then evaluate the theoretical impact of assimilating the observations. This assimilation framework was presented in Yoshimura et al. (2014) and

60 has already been applied by Toride et al. (2021) and Tada et al. (2021). Here we simulate the observations in line with the temporal and horizontal coverage achieved by the IASI (Infrared Atmospheric Sounding Interferometer, Clerbaux et al., 2009) satellite sensor. We simulate the IASI data of q, T, and δD as generated for the free troposphere by using the retrieval processor MUSICA (MUlti-platform remote Sensing of Isotopologues for investigating the Cycle of Atmospheric water, Schneider et al., 2016, 2022). We evaluate the analyses of the atmospheric fields of q, T, δD, the vertical velocity (ω), the latent heating rate

65 (Q_2), and the precipitation rate (Prcp). The latter three are strongly coupled and linked to climate feedbacks and weather events. The atmospheric dynamics (expressed among others by ω) is coupled to Q_2, which in turn affects the vertical thermal structure and thus dynamics. Prcp describes the removal of moisture from the atmosphere, which in turn affects the Q_2 and radiative heating potential.

This study is complementary to Toride et al. (2021), where observations from different platforms and different temporal and

70 spatial coverages were used (satellite, radiosonde and surface observations). The different observational techniques provide diverse information; however, using observations that have a spatial and temporal coverage that differs from the coverage of the IASI δD data, makes it difficult to understand whether an improvement in the analysis is due to the complementarity of the information provided by δD or from the complementary coverage (the coverage of IASI δD is much better than the coverage of radiosonde data and more homogeneous than the coverage of the surface data and data from geostationary satellites, see Figs. S4

75 and S5 in the Supporting Information of Toride et al., 2021). In our OSSE all observations have the same spatial and temporal coverage (the coverage of the MUSICA IASI water isotopologue satellite data, Diekmann et al., 2021a), which assures that any improvement in the analysis by an additional assimilation of δD is due to the complementarity of the information provided by δD and not affected by different coverages. Furthermore, we investigate the assimilation of δD in addition to the assimilation of satellite observations of q and T. The latter (i.e. T satellite observations) were not considered in Toride et al. (2021), despite

80 the fact that they are available with very good quality. Moreover, in addition to the general impact study given by Toride et al. (2021), this work investigates the situations when the isotopologue observations can make a unique contribution (versus the situations when they have no significant impact).

In Tada et al. (2021) real IASI δD observations (only δD observations) were assimilated and it was shown that such assimilation leads to a better agreement with the ERA5 reanalyses (Hersbach et al., 2020) than not assimilating any data. However,

Table 1. Table summarizing the typical free tropospheric observational error (σ_o) used for the assimilation experiments. The σ_o values are calculated as the root-squares-sum of MUSICA IASI retrieval noise error (σ_r) and the spatial representativeness error (σ_s, due to resampling the horizontally highly-resolved MUSICA IASI data onto the relatively coarse 200×200 km IsoGSM grid).

Observation	σ_o	σ_r	σ_s
q	0.30 g/kg	0.12 g/kg	0.27 g/kg
T	1.2 K	1.0 K	0.7 K
δD	14‰	10‰	10‰

they did not investigate the much larger impact that can already be achieved by assimilating more easily observable data like q and T. In this context, our study has a very different focus: we first estimate the impact of assimilating easily observable data (q and T) and then evaluate the impact of additionally assimilating δD observations.

The manuscript is structured as follows: section 2 describes the simulated data and the OSSE, the performed assimilation experiments, and the analysed atmospheric variables and the methods used for evaluating their quality. In Sect. 3, we give an overview on the analyses quality improvements achieved by the different assimilation experiments. Section 4 examines in detail the atmospheric events when the additional assimilation of δD has the strongest impact and it briefly discusses the possibilities of achieving such impacts also for real world analyses. A summary of the study is given in Sect. 5.

2 Data and evaluation

2.1 Data simulations

We use the isotopologue enabled atmospheric general circulation model IsoGSM (Yoshimura et al., 2008) and simulate the atmospheric state for the two months of July and August 2016, in 6 hour time steps, with a spectral model grid resolution T62 (about 200 km horizontal resolution and 28 vertical sigma levels). We use these simulations as the truth and refer to it in the following as the nature data ($x_{n_{i,j}}$, where the index i indicates the time step and the index j the location).

For our OSSE we interpolate the $x_{n_{i,j}}$ data to the locations and time step where satellite observations are available. However, MUSICA IASI data resampled to the IsoGSM grid points are used, which strongly reduces the number of observational data: while there are typically 30 high quality observations each day in a 200×200 km area (Diekmann et al., 2021a), in the OSSE used here this number is reduced to typically 1-2. We simulate observations of q, T, and δD in the middle troposphere (at about 550 hPa). The observational error variances (σ_o^2) are estimated as the sum of spatial representativeness error variance (σ_s^2) and retrieval error variance (σ_r^2). The σ_r value is the mean error estimated for the MUSICA IASI data within a IsoGSM grid box (it is typically 0.12 g/kg, 1.0 K, and 10‰ for q, T, and δD, respectively, Diekmann et al., 2021a). For the σ_s values we use the standard deviations of the MUSICA IASI data within the IsoGSM grid box, which is generally of a similar order as σ_r. Table 1 gives a summary of the typically assumed observational errors.

The different ensemble members are calculated by IsoGSM but with initialisations that are slightly different to the initialisation used from the nature run. We calculate an ensemble with 96 members, i.e. for each time step we make 96 additional
110　IsoGSM runs ($N_{\text{ens}} = 96$).

2.2 Data assimilation with a Kalman filter

For the data assimilation we use the Local Ensemble Transform Kalman Filter (LETKF, e.g. Hunt et al., 2007) method as developed for its use with water isotopologue data by Yoshimura et al. (2014). The Kalman filter based data assimilation technique optimally combines a model forecast with an observation by considering the respective model and observational
115　uncertainties (Kalman, 1960). The result is a best estimate of the atmospheric state (the analyses state vector, $\boldsymbol{x}^a$):

$$\boldsymbol{x}^a = \boldsymbol{x}^b + \mathbf{K}(\boldsymbol{y} - \mathbf{H}\boldsymbol{x}^b), \tag{2}$$

where $\boldsymbol{x}^b$ is the so-called background state (the model forecast), $\boldsymbol{y}$ the observation vector, and $\mathbf{H}$ the observational operator (a matrix operator which maps the model state into the observation space). The matrix operator $\mathbf{K}$ is the Kalman gain:

$$\begin{aligned}\mathbf{K} &= \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} \\ 120 \qquad &= (\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} + \mathbf{B}^{-1})^{-1}\mathbf{H}^T\mathbf{R}^{-1}, \end{aligned} \tag{3}$$

where the first and second line are the so-called m- and n-forms, respectively (whose equivalence is shown, for instance, in Chapt. 4 of Rodgers, 2000). The matrix $\mathbf{B}$ is the uncertainty covariance of the background state (it is calculated as the covariance of the different ensemble runs and thus captures the uncertainty of the model forecasts). Its inverse ($\mathbf{B}^{-1}$) is in the following also referred to as the background knowledge information matrix (it is a measure for the knowledge about the
125　atmospheric state including the statistical dependency of different atmospheric state variables). The matrix $\mathbf{R}$ is the uncertainty covariance of the observational state (it captures the uncertainties of the observations). If we substitute in Eq. (2) $\mathbf{K}$ by the second line of Eq. (3) and $\boldsymbol{y}$ by $\mathbf{H}\boldsymbol{x}$ (the observation $\boldsymbol{y}$ is the actual atmospheric state $\boldsymbol{x}$ mapped to the observational domain) we get:

$$\boldsymbol{x}^a = \boldsymbol{x}^b + (\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} + \mathbf{B}^{-1})^{-1}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}(\boldsymbol{x} - \boldsymbol{x}^b), \tag{4}$$

130　which reveals that the Kalman filter weights the impact of the background and the observation on the analyses reciprocally according to their respective uncertainties. More details on the used LETKF settings, like the localization, the covariance inflation or ensemble size choice, are given in Text S2 of the supplement of Toride et al. (2021).

Our assimilation experiments use observations of specific humidity (q), atmospheric temperature (T) and the isotopologue ratio of water vapour (δD) at about $550\,\text{hPa}$, which is the pressure level, where the MUSICA IASI products have generally a
135　very good quality (high sensitivity and low uncertainty). An overview on the performed different assimilation experiments is given by Table 2.

Table 2. Table summarizing the different assimilation experiments used in this study. The column "Assimilated observations" lists the observations used for the experiment, the column "Symbol" shows the symbol used in the following when referring the respective experiment, the column "Corresponding $\Delta_{i,j}$" shows the symbol used for the corresponding $\Delta_{i,j}$ value, calculated according to Eq. (6), and the column "Corresponding RMSD" shows the symbol used for the corresponding RMSD value, calculated according to Eqs. (7).

Assimilated observations	Symbol	Corresponding $\Delta_{i,j}$	Corresponding RMSD
No observations	$\{\}$	$\Delta_{i,j}\{\}$	$\mathrm{RMSD}\{\}$
q	$\{q\}$	$\Delta_{i,j}\{q\}$	$\mathrm{RMSD}\{q\}$
T	$\{T\}$	$\Delta_{i,j}\{T\}$	$\mathrm{RMSD}\{T\}$
$\delta \mathrm{D}$	$\{\delta \mathrm{D}\}$	$\Delta_{i,j}\{\delta \mathrm{D}\}$	$\mathrm{RMSD}\{\delta \mathrm{D}\}$
q and T	$\{q,T\}$	$\Delta_{i,j}\{q,T\}$	$\mathrm{RMSD}\{q,T\}$
q and $\delta \mathrm{D}$	$\{q,\delta \mathrm{D}\}$	$\Delta_{i,j}\{q,\delta \mathrm{D}\}$	$\mathrm{RMSD}\{q,\delta \mathrm{D}\}$
q, T and $\delta \mathrm{D}$	$\{q,T,\delta \mathrm{D}\}$	$\Delta_{i,j}\{q,T,\delta \mathrm{D}\}$	$\mathrm{RMSD}\{q,T,\delta \mathrm{D}\}$

2.3 Evaluation of the analyses quality

In the assimilation step the ensemble members are corrected according to the information provided by the observation. This results in an ensemble of analysed data. For convenience we interpolate the analyses fields to a regular $2.5° \times 2.5°$ horizontal grid and to 17 vertical pressure levels between 1000 and 10 hPa. We use the mean value of these analysed data (i.e. the ensemble mean values) as representative for the analysis. For a time step i and a location j this ensemble mean is

$$\bar{x}_{i,j} = \frac{1}{N_{\mathrm{ens}}} \sum_{m=1}^{N_{\mathrm{ens}}} x_{m_{i,j}}, \tag{5}$$

where $x_{m_{i,j}}$ is the ensemble member m at time step i and location j. For each location and time step (each event) we calculate the difference of the ensemble mean and the nature data.

$$\Delta_{i,j} = \bar{x}_{i,j} - x_{n_{i,j}}. \tag{6}$$

This $\Delta_{i,j}$ captures the uncertainty for the single event corresponding to time step i and location j. This is what we want to evaluate.

We then calculate the root-mean-squares of the $\Delta_{i,j}$ uncertainties (root-mean-squares-differences, RMSD) for all events belonging to a group of events A:

$$\mathrm{RMSD} = \sqrt{\sum_{(i,j)\in A} \Delta_{i,j}^2 \bigg/ \sum_{(i,j)\in A} 1}. \tag{7}$$

The group of events A can include all events (sum over all time steps and locations) or only selected events that fulfil certain criteria. The RMSD values is a statistically robust metric representing the uncertainty of the analyses data for the events that belong to the group of events A.

Table 3. Table with skill values discussed in this study. The column "Description of skill" outlines what assimilation experiments are used for calculating the skills (the evaluated experiment and the reference experiment, with respect to which the evaluation is performed), the column "Symbol" shows the symbol used in the text when referring the respective skill, and the column "Skill calculation" describes how the skill is calculated according to Eq. (8).

Description of skill	Symbol	Skill calculation
$\{q\}$ wrt $\{\}$	$\{q\}_{\{\}}$ skill	$\frac{\mathrm{RMSD}\{\}-\mathrm{RMSD}\{q\}}{\mathrm{RMSD}\{\}}$
$\{T\}$ wrt $\{\}$	$\{T\}_{\{\}}$ skill	$\frac{\mathrm{RMSD}\{\}-\mathrm{RMSD}\{T\}}{\mathrm{RMSD}\{\}}$
$\{\delta\mathrm{D}\}$ wrt $\{\}$	$\{\delta\mathrm{D}\}_{\{\}}$ skill	$\frac{\mathrm{RMSD}\{\}-\mathrm{RMSD}\{\delta\mathrm{D}\}}{\mathrm{RMSD}\{\}}$
$\{q,T\}$ wrt $\{\}$	$\{q,T\}_{\{\}}$ skill	$\frac{\mathrm{RMSD}\{\}-\mathrm{RMSD}\{q,T\}}{\mathrm{RMSD}\{\}}$
$\{q,T\}$ wrt $\{q\}$	$\{q,T\}_{\{q\}}$ skill	$\frac{\mathrm{RMSD}\{q\}-\mathrm{RMSD}\{q,T\}}{\mathrm{RMSD}\{q\}}$
$\{q,\delta\mathrm{D}\}$ wrt $\{q\}$	$\{q,\delta\mathrm{D}\}_{\{q\}}$ skill	$\frac{\mathrm{RMSD}\{q\}-\mathrm{RMSD}\{q,\delta\mathrm{D}\}}{\mathrm{RMSD}\{q\}}$
$\{q,T,\delta\mathrm{D}\}$ wrt $\{q,T\}$	$\{q,T,\delta\mathrm{D}\}_{\{q,T\}}$ skill	$\frac{\mathrm{RMSD}\{q,T\}-\mathrm{RMSD}\{q,T,\delta\mathrm{D}\}}{\mathrm{RMSD}\{q,T\}}$

From the RMSD values we then detemine the skill of an assimilation experiment as:

$$\mathrm{Skill} = \frac{\mathrm{RMSD}\{\mathrm{ref}\} - \mathrm{RMSD}\{\mathrm{exp}\}}{\mathrm{RMSD}\{\mathrm{ref}\}}, \tag{8}$$

where $\mathrm{RMSD}\{\mathrm{exp}\}$ corresponds to the experiment we want to evaluate and $\mathrm{RMSD}\{\mathrm{ref}\}$ to the reference experiment with respect to which we want do do the evaluation. So the skill informs about the relative change of the RMSD value obtained from an assimilation experiment with respect to a reference assimilation experiment, i.e. it documents the relative improvement of the analyses data with respect to a reference. We use this skill value throughout the paper for evaluating the quality of the analyses obtained by the different assimilation experiments. Table 3 gives an overview on the different kind of skill values that are discussed in this study.

For this study, we consider the analyses of the variables specific humidity (q), atmospheric temperature (T), and water vapour isotopologue ratio ($\delta\mathrm{D}$). For these variables we have also observations that are assimilated (see Sect. 2.2). In addition, we evaluate the analyses of variables that describe the water cycle, but we do not assimilate observations of these variables. The analysed atmospheric water cycle variables are vertical velocity (ω), latent heating rate (Q_2) and precipitation rate (Prcp). The Q_2 values are calculated according to the budget analysis (Yanai et al., 1973):

$$Q_2 = -L\left(\frac{\partial q}{\partial t} + \boldsymbol{v} \cdot \nabla q + \omega \frac{\partial q}{\partial p}\right), \tag{9}$$

where L is the latent heat of net condensation, q is the specific humidity, $\boldsymbol{v}$ is the horizontal wind vector, ω is the vertical velocity, and p is the pressure.

We work with daily mean analyses data (for the $30°$S - $30°$N region) of July and August 2016. The ensemble simulations are made using 96 different initial conditions. Therefore, the ensemble mean at the beginning of the simulation period represents climatology. The first three weeks of the simulation (beginning of July) is a "spin-up" period, when the analyses gradually

approximate the nature data by assimilating enough observations. In order to avoid impacts of this "spin-up" period, respective data are excluded from the the evaluation study, which is then made for the mid-July to end of August period (covering 40 days).

3   Overview on assimilation impacts

This section gives an overview on the assimilation impacts. For this purpose we calculate the RMSD values using all events (averaging is performed over all time steps and locations). Equation (7) can then be written as:

$$\text{RMSD} \quad = \quad \sqrt{\frac{1}{N_{\text{loc}}N_{\text{tim}}} \sum_{j=1}^{N_{\text{loc}}} \sum_{i=1}^{N_{\text{tim}}} \Delta_{i,j}^2}, \tag{10}$$

where N_{loc} is the number of all locations (here we investigate the 30°S-30°N region with a $2.5° \times 2.5°$ resolution, i.e. $N_{\text{loc}} = 3600$), and N_{tim} is the number of all time steps (here we work with 24 h time steps covering 40 days, i.e. $N_{\text{tim}} = 40$).

Because for this calculation data of continuous time series are used, we cannot assume independence of the different data when estimating the uncertainty of the RMSD values. For this reason, we use the circular block bootstrap method (e.g. Wilks, 2019) for the RMSD uncertainty estimation (the method is also explained in the supplement of Toride et al., 2021). We resample these data 10000 times, which provides a representative distribution of possible RMSD values. Here we use the half of the difference between the respective 15.9th and 84.1th percentile estimates as the 1σ uncertainty of the RMSD value, which we then propagate to the skill values.

3.1   Standard observations and single type of observations

Observations of free tropospheric q and T contain important information on the atmospheric state (among others on the water cycle variables ω, Q_2, and Prcp) and are available as standard products from different satellite data processors at global scale, daily coverage, and with good precision. We use the experiments that assimilate these observations in order to understand which level of analyses quality can be already achieved by assimilating the easily observable variables.

First we calculate the skills achieved when assimilating q observations only using no data assimilation as the reference, i.e. here RMSD_{ref} of Eq. (8) is for ensemble means ($\bar{x}_{i,j}$) obtained without assimilating any observation (no data assimilation step). This skill is referred to in the following as the $\{q\}_{\{\}}$ skill (see Table 3). The light blue lines in Figure 1a-e show the vertical dependency of the $\{q\}_{\{\}}$ skills (for Prcp there is naturally no vertical dependency, Fig. 1f). The area around the lines indicates the 1σ uncertainty of the skills. Because we assimilate the observations of q at about 550 hPa, highest skill values are generally achieved in the free troposphere around 550 hPa. The dotted lines correspond to the pressure levels at 775 and 350 hPa, which delimits the vertical range we use as representative for the free troposphere and for which we perform dedicated evaluations in Sect.4.

In a second experiment we assimilate observations of q together with T, which comes very close to an assimilation of relative humidity data. For the evaluation we again calculate the skills with respect to no data assimilation (in the following referred to as the $\{q,T\}_{\{\}}$ skill, see Table 3). The black lines in Figs. 1a-f give an overview on the achieved $\{q,T\}_{\{\}}$ skills. Compared to the $\{q\}_{\{\}}$ skills, these skills are larger in particular for T around 550 hPa (Fig. 1b), because T at 550 hPa is the additional
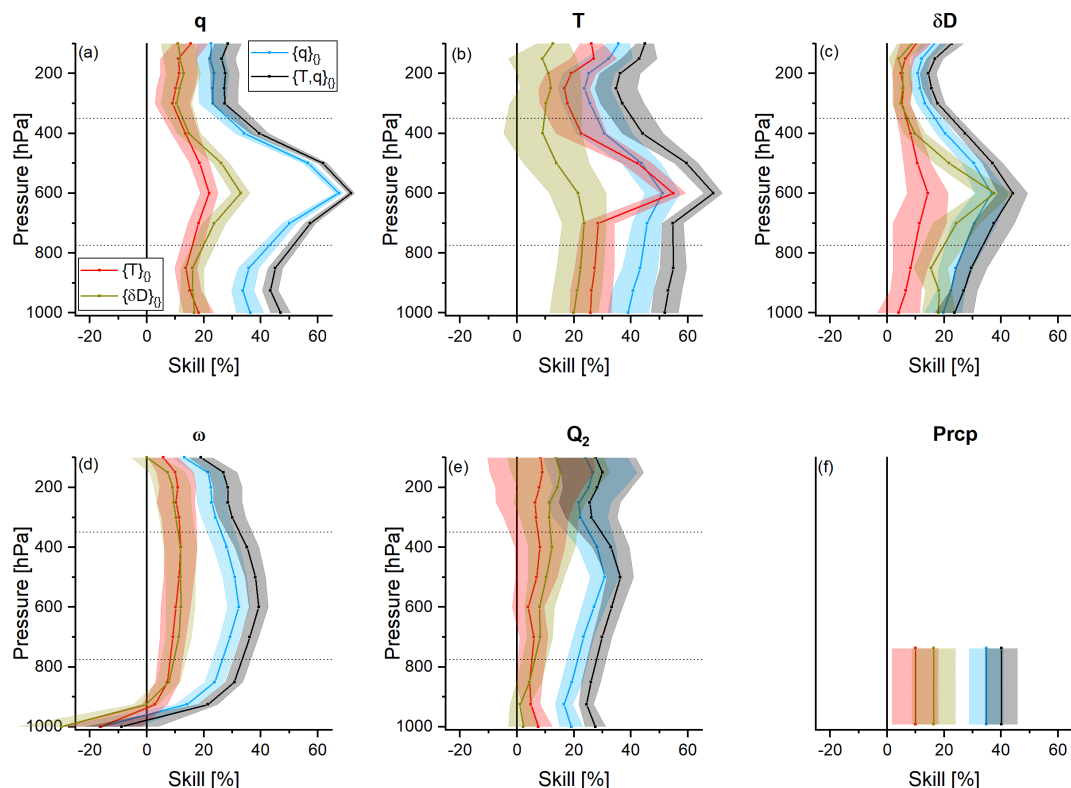
Figure 1. Vertical profiles of the skills achieved by assimilating the standard observations (only q or q together with T) or only one type of observation (only T or only δD) versus assimilating no observations. Light blue: $\{q\}_{\{\}}$ skill; Black/grey: $\{q,T\}_{\{\}}$ skill; red: $\{T\}_{\{\}}$ skill; dark yellow: $\{\delta D\}_{\{\}}$ skill. The area around the lines represents the 1σ uncertainty. (a) For specific humidity (q), (b) for temperature (T), (c) for the isotopologue ratio (δD), (d) for vertical velocity (ω), (e) skill for latent heating (Q_2), and (f) for precipitation (Prcp).

observation we assimilate. The additional assimilation of T has also significantly positive impacts on q and Q_2 above $700\,\mathrm{hPa}$ and on δD between 500 and $300\,\mathrm{hPa}$.

By assimilating the standard observations q and T, we achieve skills of up to 60% for q and T around $600\,\mathrm{hPa}$. Also for the other variables (δD, ω, Q_2, and Prcp) – for which no respective observations are assimilated – we get skills that are often between between 30% and 40%.

In a third and fourth experiment we test the impact when assimilating only T or only δD observations. For the evaluation we again calculate the skills with respect to no data assimilation (in the following referred to as the $\{T\}_{\{\}}$ or $\{\delta D\}_{\{\}}$ skills, respectively, see Table 3). The red lines in Figs. 1a-f give an overview on the achieved $\{T\}_{\{\}}$ skills. The values are up to 20% for q (Fig. 1a) and up to 10% for δD, ω, Q_2, and Prcp (Figs. 1c-f), which is significantly smaller than the $\{q\}_{\{\}}$ skill values. Even for T around $550\,\mathrm{hPa}$, the $\{T\}_{\{\}}$ skills are only slightly larger than the $\{q\}_{\{\}}$ skills (Fig. 1b), i.e. even for the pressure level where the observations are assimilated, assimilating q observations has almost the same positive impact on the T analyses as assimilating T observations. For the $\{\delta D\}_{\{\}}$ skills (dark yellow lines in Figs. 1a-f) the situation is similar. For q, ω, Q_2 and

9

Prcp, the skills are of a similar order as the $\{T\}_{\{\}}$ skills and significantly smaller than the $\{q\}_{\{\}}$ skills (Figs. 1a, d-f). For the T analyses the $\{\delta\text{D}\}_{\{\}}$ skills are smaller than the $\{T\}_{\{\}}$ skills (Fig. 1b) and for the δD analyses it is vice versa (Fig. 1c). For the δD analyses around 550 hPa, the $\{\delta\text{D}\}_{\{\}}$ skill is close to 40%, which is, however, only slightly larger than the respective $\{q\}_{\{\}}$ skill.

220 ## 3.2   Complementarity of additional observations

The previous subsection reveals that already assimilating the standard observations (q and T) strongly improves the analyses and that the individual impact of q observations is stronger than the individual impact of T and δD observations (the latter two have a similar impact). In this subsection, we investigate the complementarity of assimilating T and δD observations on top of q observations and of assimilating δD observations on top of q and T observations. Our particular interest is in the

225 complementarity of δD, whose observations have become only recently available on global scale and daily coverage. In order to evaluate the complementarity, we calculate the additional skills using the assimilation of the standard observations (q or q together with T) as the reference, i.e. RMSD_{ref} of Eq. (8) is for ensemble means ($\bar{x}_{i,j}$) obtained when assimilating observation of q or when assimilating observations of q together with T.

Figure 2a-f shows the same as Fig. 1, but for the additional skills. In order to document the values achieved typically for

230 additional skills, we first evaluate the additional skills for assimilating observations of T in addition to observations of q only, i.e. we calculate the skill of assimilating observations of q together with T using the assimilation of q observations only as the reference (in the following referred to as the $\{q,T\}_{\{q\}}$ skill, see Table 3). The $\{q,T\}_{\{q\}}$ skills are represented by the magenta line (and the shaded area is the 1σ uncertainty of the skill). We observe that, when adding T as an additional observation, the additional skill is highest for T and close to 40% at 500-600 hPa; for other vertical pressure levels it is smaller, but still above

235 or close to 20% (Fig. 2b). Additional observations of T also improve the analyses of q, δD, Q_2, ω and Prcp (although less than for T): the $\{q,T\}_{\{q\}}$ skills are close to 15% above 800 hPa for q (Fig. 2a) and close to 10% around 600 hPa for δD and ω (Figs. 2c+d), above 500 hPa for Q_2 (Fig. 2e), and for Prcp (Figs. 2f).

The green line in Fig. 2 represents the additional skills for assimilating observations of δD in addition to observations of q, i.e. the skill is calculated for assimilating q and δD observations using the assimilation of q observations only as the

240 reference (in the following referred to as the $\{q,\delta\text{D}\}_{\{q\}}$ skill, see Table 3). For q, we see a significant skill of about 10% for pressure levels above 500 hPa and for T, a significant positive skill is only achieved for the 600 hPa pressure level (Fig. 2a+b, with significant we mean here that the calculated skill value is larger than the estimated 1σ uncertainty of the skill, which is represented by the shaded area around the green line). For δD, the additional skill is clearly significant and up to 25% around 600 hPa, and typically above 10% for other the pressure levels above 400 hPa (Fig. 2c), which is reasonable because the

245 additional observation of δD is well suited for improving the quality of the δD analyses. For ω and Q_2, the $\{q,\delta\text{D}\}_{\{q\}}$ skills are significant over a large part of the free troposphere, but generally smaller than the respective $\{q,T\}_{\{q\}}$ skills (Figs. 2d+e). For Prcp, the $\{q,\delta\text{D}\}_{\{q\}}$ and $\{q,T\}_{\{q\}}$ are very similar and close to 10% (Fig. 2f).

In a final assimilation experiment we investigate the additional skills for assimilating observations of δD in addition to observations of q and T, i.e. the skill is calculated for assimilating q, T, and δD observations using the assimilation of q and
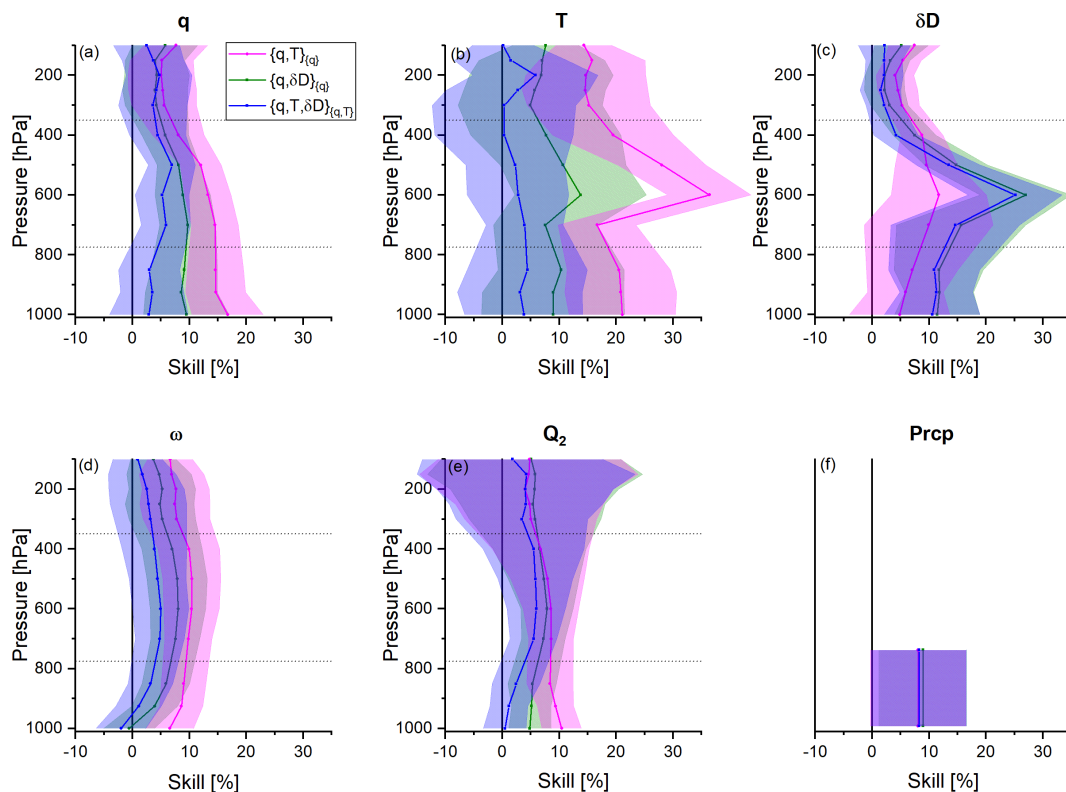
Figure 2. Same as Fig. 1, but for the skills achieved by assimilating observations on top of the standard observations (q and T): Magenta: the $\{q,T\}_{\{q\}}$ skill; olive: $\{q,\delta D\}_{\{q\}}$ skill; blue: $\{q,T,\delta D\}_{\{q,T\}}$ skill.

250 T observations as the reference (in the following referred to as the $\{q,T,\delta D\}_{\{q,T\}}$ skill, see Table 3). The overview for the respective additional skills is shown as blue lines in Figs. 2a-f. Concerning the analyses of T, the additional assimilation of δD observations cannot further improve the analyses quality as already achieved by assimilating observations q and T, i.e. the respective $\{q,T,\delta D\}_{\{q,T\}}$ skills are not significant (Fig. 2b). For the q analyses, the $\{q,T,\delta D\}_{\{q,T\}}$ are weakly positive and significant between 400 and 800 hPa (skills are generally below 5%, Fig. 2a). For δD, the $\{q,T,\delta D\}_{\{q,T\}}$ skills are as large

255 as the $\{q,\delta D\}_{\{q\}}$ skills (up to 25% around 600 hPa, Fig. 2c), i.e. the δD observations are much more important than the T observations for improving the analysis quality of δD. For ω and Q_2, there are weak significant skills of about 5% between 600 and 700 hPa (Fig. 2d+e) and for Prcp, the skill is close the 10% and weakly significant (Fig. 2f). This means that δD observations provide information on top of q and T observations, that is useful for further improving the quality of the ω, Q_2, and Prcp analyses.

4 The complementarity of δD

The overview study of the previous section reveals a positive (although weak) impact due to the additional assimilation of δD observations. Theoretically the δD data contain unique information on phase transitions, i.e. it might be expected that δD observations can in particular improve the quality of the analyses for atmospheric conditions that involve strong and/or repeated cycles of condensation (or evaporation) processes. In this section we examine the adequacy of this hypothesis in more detail and focus on the analyses of data averaged over a free tropospheric pressure range (775-350 hPa, indicated by the dotted lines in Figs. 1 and 2).

4.1 Analyses quality and atmospheric latent heating

If water condenses latent heat is released to the atmosphere and for the evaporation of water energy is needed (heat is removed from the atmosphere, i.e. negative latent heating). For this reason we can use the latent heating rate (Q_2) as a proxy for phase transitions.

Figure 3 depicts the dependency of the free tropospheric analyses errors (the $\Delta_{i,j}$ values, see Eq. (6)) on the free tropospheric latent heating rate (Q_2 as simulated by the nature run). As in Figs. 1 and 2 we investigate the analyses of the atmospheric variables q, T, δD, ω, Q_2, and Prcp. We examine the low latitudes (30°S - 30°N, with a $2.5° \times 2.5°$ horizontal resolution) for 40 days, i.e. in total we have 144000 events. In order to visualize the distribution of these large amount of data points, we calculate the data densities as follows: we generate 60 equidistant Q_2-bins covering all occurring Q_2. Then we calculate the density distribution of the $\Delta_{i,j}$ values in each Q_2-bin and sum up the number of data points belonging to the highest $\Delta_{i,j}$ densities until we consider 98% of all the $\Delta_{i,j}$ values occurring for the considered Q_2-bin. These 98% areas are depicted in Fig. 3. The grey filled area represents the $\Delta\{\}$ distribution (i.e. for the $\Delta_{i,j}$ values when no observations are assimilated), and the magenta and blue lines the 98% contour lines for the $\Delta\{q,T\}$ and $\Delta\{q,T,\delta\text{D}\}$ distributions (i.e. for the $\Delta_{i,j}$ values achieved when we assimilate q and T observations and q, T and δD observations, respectively).

When no observations are assimilated, the latent heating events can hardly be identified and the Q_2 uncertainty is almost as large as the actual Q_2 value (the $\Delta\{\}$ distribution in Fig. 3e aligns very closely with the black dashed anti-diagonal). The $\Delta\{q,T\}$ and $\Delta\{q,T,\delta\text{D}\}$ distributions dissipate from the anti-diagonal and approximate better the $\Delta(Q_2)$-zero line. This reduction of the Q_2 uncertainty is significantly more pronounced for the $\Delta\{q,T,\delta\text{D}\}$ than for the $\Delta\{q,T\}$ distributions, and largest for high actual Q_2 values. However, despite the significant correction, the uncertainty is still largest for the highest Q_2 values. This means that, although the events of strong latent heating can be much better identified by assimilating the observations, the absolute strength of the heating is still underestimated.

For the analyses of vertical velocity (ω, Fig. 3d) and precipitation rate (Prcp, Fig. 3f) the results are very similar to those of Q_2: without assimilating any data ($\Delta\{\}$ densities), the analyses are very uncertain for high heating rates (distribution of $\Delta_{i,j}$ values is far away from the respective Δ-zero lines). Actually vertical velocity, precipitation rate, and heating rate are strongly correlated, which means that the events with strong upward motion of air and/or with high precipitation rates are almost not identified. By assimilating q and T observations, this uncertainties can be strongly reduced. A further significant reduction
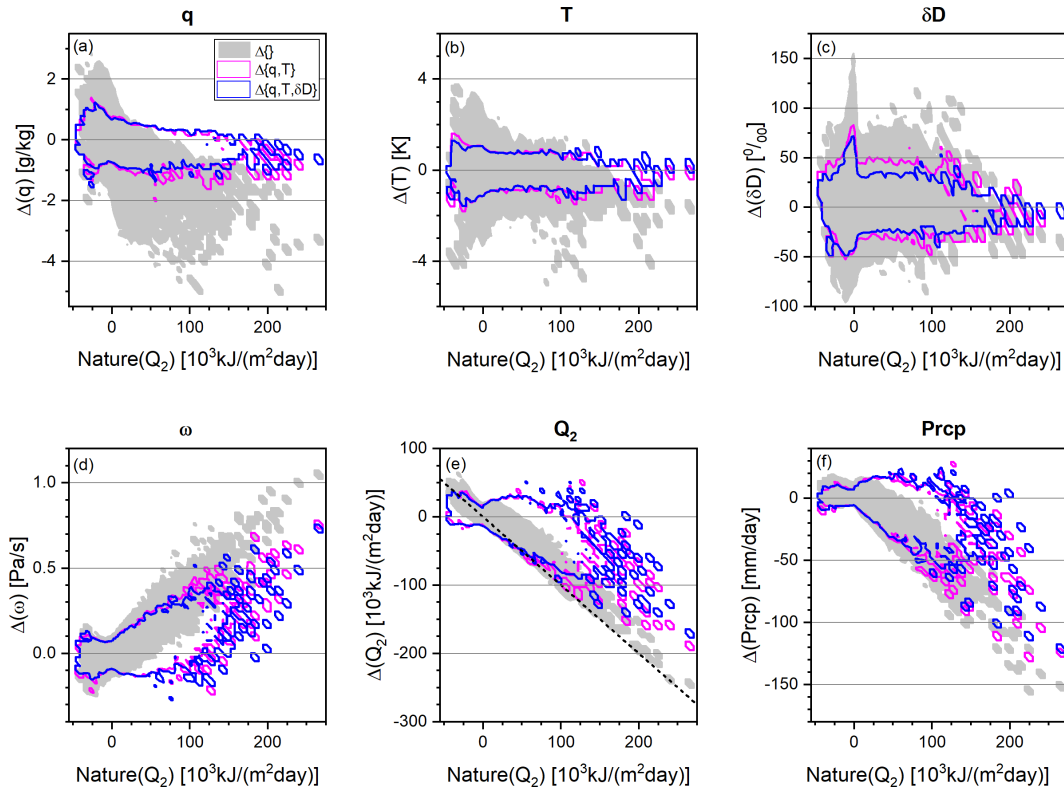
12

Figure 3. Dependency of the analyses errors on the free tropospheric heating rate ($\mathrm{Nature}(Q_2)$). Shown are the areas that contain 98% of all the $\Delta_{i,j}$ values for a given $\mathrm{Nature}(Q_2)$ value. Grey: 98% area for no data assimilation (the $\Delta\{\}_{i,j}$ data distribution). Magenta line: 98% contour line for the $\Delta\{q,T\}_{i,j}$ data distribution. Blue line: 98% contour line for the $\Delta\{q,T,\delta\mathrm{D}\}_{i,j}$ data distribution. (a) for specific humidity (q), (b) for temperature (T), (c) for the isotopologue ratio ($\delta\mathrm{D}$), (d) for vertical velocity (ω), (e) skill for latent heating (Q_2), and (f) for precipitation (Prcp, e).

(in particular for events when the uncertainty is very high) can be achieved by assimilating $\delta\mathrm{D}$ observations in addition to the observations of q and T.

295 Figures 3a-c show the distributions of the $\Delta_{i,j}$ values for the variables q, T, and $\delta\mathrm{D}$. There are only weak correlations between the analyses uncertainties and the $\mathrm{Nature}(Q_2)$ data. Although the $\Delta(q)_{i,j}$ values tend to be positive for negative heating rates and negative for positive heating rates, this dependency is much weaker as what is observed for ω, Q_2, and Prcp. A rather weak dependency is also observed in $\Delta(T)_{i,j}$ and to an even smaller extent in $\Delta(\delta\mathrm{D})_{i,j}$: in both cases the largest positive or negative values occur for negative Q_2. The assimilation of q and T observations approximates the $\Delta_{i,j}$ values

300 to the respective Δ-zero lines. Concerning the q analyses the additional assimilation of $\delta\mathrm{D}$ observations further reduced the uncertainty in particular for events with high latent heating rates (for high $\mathrm{Nature}(Q_2)$ the blue contour line better approximates the Δ-zero line than the magenta contour line, Fig. 3a). For T the additional impact when assimilating $\delta\mathrm{D}$ observation seems also to be slightly larger for high positive or negative heating rates (Fig. 3b). For $\delta\mathrm{D}$ an the additional impact of assimilating
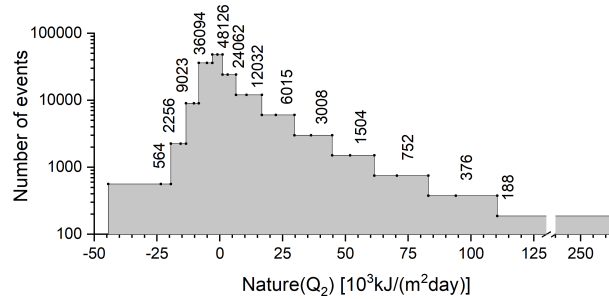
Figure 4. Abundance chart showing the number of events for each of the 13 Q_2-bins used for classifying different latent heating strengths.

δD observations seems to be largely independent on the heating rates (blue contour lines better approximates the Δ-zero line than the magenta contour for all Q_2 values, Fig. 3c).

4.2 The unique δD assimilation impact

In order to better quantitatively document the dependency between the latent heating rate and the additional impact of assimilating δD observations, we investigate in this subsection how the skills depend on the Q_2 nature data.

Figure 4 shows the abundances of events corresponding to 13 different Q_2 heating rate bins. We have the highest abundances for heating rates that are close to zero. The three bins corresponding to Nature(Q_2) values between -8.3 and $+6.5 \times 10^3$ kJ/(m^{2}day) comprise together 108282 events, which is $\frac{108282}{144000} = 75.20\%$ of all events. Extreme heating rates are rare, e.g. the bin corresponding to Nature(Q_2) values between $+110.5$ and $+262.5 \times 10^3$ kJ/(m^{2}day) only comprises 188 events, i.e. only $\frac{188}{144000} = 0.13\%$ of all the data points. However, these extreme events are responsible for almost all the intense precipitation events, and it is very important to improve the analyses in a way that allows a better identification of these extreme events.

We use the binning from Fig. 4 for evaluating the skill dependence on Q_2. Therefore, we calculate the RMSD values according to Eq. (7) for 13 different groups of events A. Each of the groups comprises the events showing heating rates as defined by the 13 different bins of Fig. 4. Then for each of the bins we calculate the skill according to Eq. (8). Because the events with high Q_2 values are generally individual events occurring on a single day, the respective groups of events do not consist of continuous time series. For the error estimation we thus assume that the events of a certain Q_2 group are independent (in the circular block bootstrap method the block size is reduced to only one event).

Figure 5 depicts the skills obtained for the 13 Nature(Q_2)-bins. The colours are as in Fig. 2. Concerning the additional skill for assimilating T observations on top of q observations ($\{q, T\}_{\{q\}}$ skill, represented by magenta colour), we cannot identify a clear dependency on the Q_2 values. The skills are more or less the same as already shown in Fig. 2 for the 775 to 350 hPa altitude range (overview examination of all events together): highest skills of between 20% and 30% are observed for T (Fig. 5b), skills between 10% and 15% are seen for q (Fig. 5a), and for δD, ω, Q_2, and Prcp the skills are generally below 10% (Fig. 5c-f).
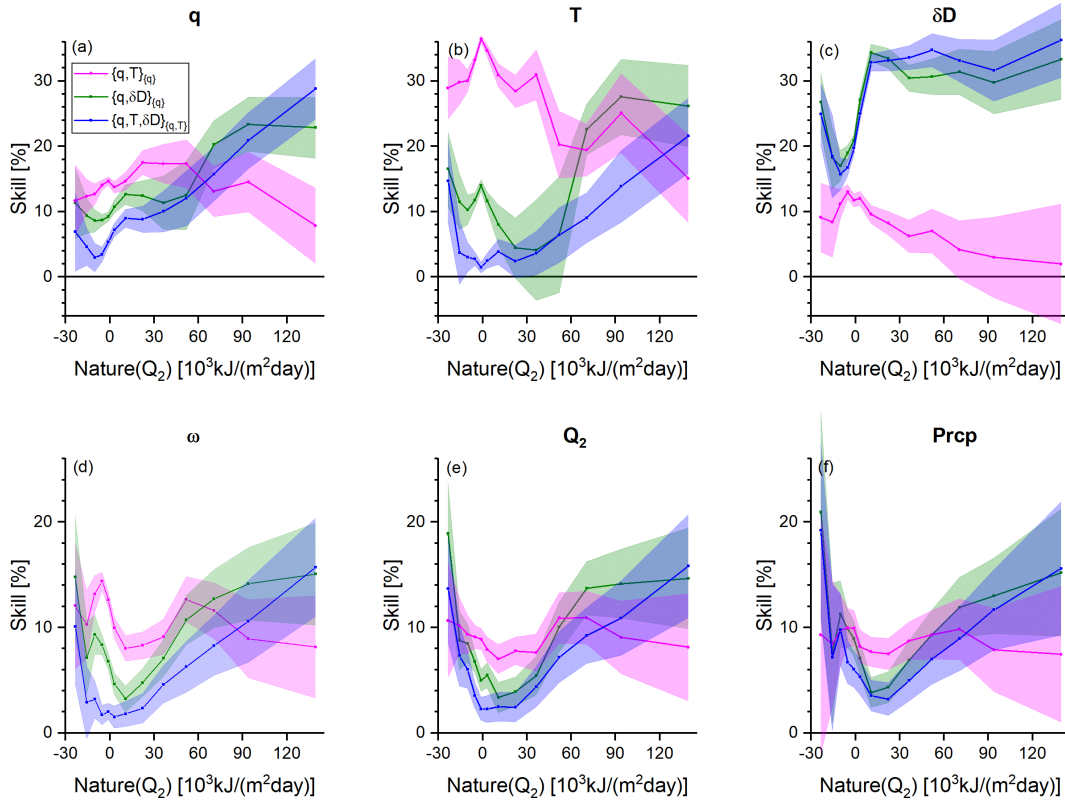
Figure 5. Dependency of the skills on the free tropospheric latent heating rate ($\mathrm{Nature}(Q_2)$). The skills are calculated for the Q_2-bins as depicted in Fig. 4. Shown are the skills achieved by assimilation observations on top of the standard observations: Magenta: the $\{q,T\}_{\{q\}}$ skill; olive: $\{q,\delta\mathrm{D}\}_{\{q\}}$ skill; blue: $\{q,T,\delta\mathrm{D}\}_{\{q,T\}}$ skill. The areas around the lines indicate the 1σ uncertainties of the skills. (a) For specific humidity (q), (b) for temperature (T), (c) for the isotopologue ratio ($\delta\mathrm{D}$), (d) for vertical velocity (ω), (e) for latent heating (Q_2), and (f) for precipitation (Prcp).

The $\{q,\delta\mathrm{D}\}_{\{q\}}$ skills (additional skills for assimilating $\delta\mathrm{D}$ observations on top of q observations, represented by olive colour) show a clear dependency on the Q_2 values. For q, T, ω, Q_2, and Prcp (Figs. 5a,b,d-f) and for low heating rates (Nature(Q_2)

330 between -15.5 and $+36.0 \times 10^3$ kJ/(m^{2}day), these skills are smaller than the respective $\{q,T\}_{\{q\}}$ skills. The respective Q_2 bins comprise 97% of all examined events, which explains that this finding is similar to the one from the overview plot (Fig. 2), except for Prcp where in the overview plot there is no clear difference between the $\{q,\delta\mathrm{D}\}_{\{q\}}$ and $\{q,T\}_{\{q\}}$ skills. However, for the rare occasions of high heating rates (negative or positive, Q_2 below -20×10^3 kJ/(m^{2}day) or above $+70\times10^3$ kJ/(m^{2}day)) the $\{q,\delta\mathrm{D}\}_{\{q\}}$ skills get as large or even larger than the $\{q,T\}_{\{q\}}$ skills. For $\delta\mathrm{D}$ (Fig. 5c) the $\{q,\delta\mathrm{D}\}_{\{q\}}$ skills are for all Q_2

335 values larger than the $\{q,T\}_{\{q\}}$ skills, naturally because the assimilation of $\delta\mathrm{D}$ observations is most important for improving the $\delta\mathrm{D}$ analyses.
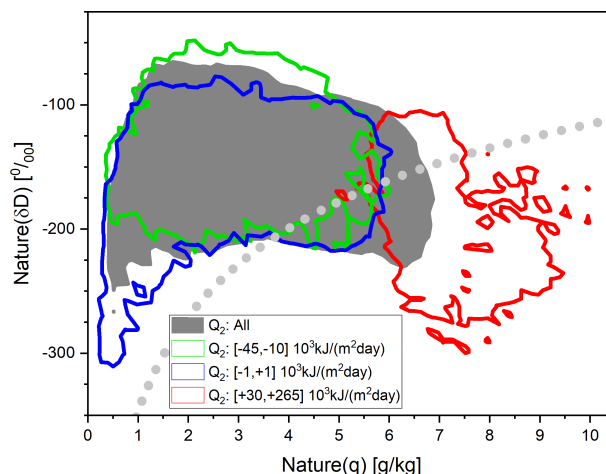
Figure 6. Different $\{q, \delta\mathrm{D}\}$-pair distributions as seen in the data from the nature run. Shown are the areas that contain 80% of all the data. Dark grey: 80% area when considering all data. Green line: 80% contour line when considering data with $\mathrm{Nature}(Q_2) < -10 \times 10^3 \, \mathrm{kJ/(m^2 day)}$. Blue line: 80% contour line when considering data with $\mathrm{Nature}(Q_2) > -1 \times 10^3 \, \mathrm{kJ/(m^2 day)}$ and $\mathrm{Nature}(Q_2) < +1 \times 10^3 \, \mathrm{kJ/(m^2 day)}$. Red line: 80% contour line when considering data with $\mathrm{Nature}(Q_2) > +10 \times 10^3 \, \mathrm{kJ/(m^2 day)}$. The dotted grey line is a typical tropical Rayleigh line, assuming the following atmospheric condition over the ocean source location: $T = 25°\mathrm{C}$, $\mathrm{RH} = 80\%$, and $\delta\mathrm{D} = -80‰$.

The $\{q, T, \delta\mathrm{D}\}_{\{q,T\}}$ skills (additional skills for assimilating $\delta\mathrm{D}$ observations on top of q and T observations, represented by blue colour) show a clear dependency on the Q_2 values. For the analyses of all the variables, the skills are highest for strongly positive or strongly negative $\mathrm{Nature}(Q_2)$ values. For $\mathrm{Nature}(Q_2)$ close to zero, the skills are very small (often below 340 5%, except for $\delta\mathrm{D}$ where it is never below 20%). Because for the large majority of events the heating rates are close to zero (see discussion above), the overview study reveals generally small $\{q, T, \delta\mathrm{D}\}_{\{q,T\}}$ skill values. However, $\delta\mathrm{D}$ observations become very important for events corresponding to extreme heating rates. Then assimilating the $\delta\mathrm{D}$ observations on top of q and T observations can strongly improve the quality of the analyses of all variables (skills of 15% and above). This is in particularly important for the analyses of ω, Q_2, and Prcp, because these variables are relatively poorly constrained by the q 345 and T observations. The most extreme ω, Q_2, and Prcp values are relatively poorly identified by assimilating only q and T observations; a better identification of these events is achieved by the additional assimilating of $\delta\mathrm{D}$ (compare the magenta and blue contour lines in Fig. 3d-f).

4.3 Discussion and outlook

The $\{q, \delta\mathrm{D}\}$-pair distributions can give valuable insight into the dominating atmospheric processes (mixing, shallow cloud 350 formation and rain-out, convection and extreme precipitation events, Noone, 2012). Figure 6 shows different $\{q, \delta\mathrm{D}\}$-pair distributions as obtained from the nature simulation run. Shown are the areas where the $\{q, \delta\mathrm{D}\}$-pairs have the highest densities

and sum up to 80% of all the available data. The distributions are determined for all 144000 events and for three different subsets of events belonging to three different latent heating categories.

The dark grey area comprises the area where 80% of all the 144000 data points are located. The thick grey dotted line

355   represents a typical tropical Rayleigh line (starting conditions: $T = 25°$C, RH $= 80\%$, and δD $= -80‰$.). We observe that for most of the events the $\{q, \delta$D$\}$-pairs are located above the Rayleigh line, indicating that mixing processes very importantly affect the nature simulations (Noone et al., 2011; González et al., 2016).

The green contour line comprises 80% of all the events with strongly negative latent heating rates (Nature(Q_2) below -10×10^3 kJ/(m^{2}day)). These subset of events is more affected by mixing between dry/depleted and humid/enriched water

360   masses than the average of all events (indicated by a $\{q, \delta$D$\}$-pair distribution moving to lower q, but higher δD values). For the subset that comprises the events where Q_2 is negligible (within $\pm 1 \times 10^3$ kJ/(m^{2}day), blue contour line), mixing between dry/depleted and humid/enriched end members is also important, but in difference to the green contour lines, here we have a significant number of $\{q, \delta$D$\}$-pairs with very low q and low δD. This strong depletion together with low humidity suggest that the dry/depleted end member is dominating the mixed water mass. We found that this is correlated to the events with the highest

365   positive vertical velocity values, i.e. strong downwelling of upper tropospheric air into the middle troposphere. This typically happens in the subtropics and makes the free tropospheric air very dry (very low relative humidities), i.e. neither condensation nor evaporation is happening, and there is no latent heating.

The red contour line comprises the events with strongly positive latent heating rates (Nature(Q_2) above $+30 \times 10^3$ kJ/(m^{2}day)). The $\{q, \delta$D$\}$-pair corresponding to these events are generally located below the Rayleigh line in the super-Rayleigh domain

370   (Noone, 2012). This strong depletion together with high humidity is caused by recurring evaporation and condensation, i.e. the same water mass experiences several condensation/evaporation processes in the atmosphere. This is a typical $\{q, \delta$D$\}$-pair signal for convective activity (e.g. Blossey et al., 2010; Diekmann et al., 2021b). According to Figs. 3 and 5, it is the δD observations made in this super-Rayleigh domain, that most importantly improve the analyses.

These findings enable us to give some outlook on the possibility of improving real world analyses with real world $\{q, \delta$D$\}$-

375   pair satellite data, similarly to the improvements as demonstrated here theoretically with our OSSE. In the MUSICA IASI $\{q, \delta$D$\}$-pair satellite data super-Rayleigh conditions are often observed. They are, for instance, reported for the central Pacific in Fig. 11 of Schneider et al. (2017) or for West Africa in Fig. 14 of Diekmann et al. (2021a) and in Chapt. 6.3 of Diekmann (2021). They are also reported for the isotopologue data generated from measurements of the TES (Tropospheric Emission Spectrometer) instrument (e.g. in Sect. 4 of Noone, 2012). This demonstrates, that current state-of-the-art infrared satellite

380   instruments like IASI or TES can deliver water vapour isotopologue data for those conditions, under which they are most impacting on the analyses. The IASI instrument and thus the MUSICA IASI data set is in particular promising in this context: measurements of IASI (or IASI-NG, the successor instrument of IASI) offer a very high horizontal and spatial coverage, and are guaranteed at least for the next two decades in the framework of the Metop and Metop-SG missions of EUMETSAT (European Organisation for the Exploitation of Meteorological Satellites, https://www.eumetsat.int/our-satellites/metop-series).

385   In addition to the availability of observations linked to the most impacting convective conditions, we need atmospheric isotopologue enabled models that capture as much as possible details of a convective atmosphere. The here used model IsoGSM

(with 200×200 km horizontal resolution) does generally a good job, which has been demonstrated in different model validation studies (e.g. Yoshimura et al., 2008; Schneider et al., 2010). However, a higher horizontal resolution and a convection permitting model setup (instead of parametrising convection as in IsoGSM) might further improve the capability of a model for correctly

390   capturing the real world multi-scale impact of convective events (e.g. Pante and Knippertz, 2019). In this context, the ongoing development of including water isotopologue simulations into different highly resolving models also used for operational weather forecasting (e.g. Pfahl et al., 2012; Eckstein et al., 2018) is very encouraging. In future work the LETKF assimilation approach could be used together with a model resolution that comes close to the horizontal resolution of the MUSICA IASI data (10-20 km). Then the model might capture many details of the convective atmosphere and moreover, all MUSICA IASI

395   observations could be used (no averaging/aggregation onto a coarser model grid would be required). On the other hand, using an Ensemble Kalman Filter together with a highly-resolving model is computationally very expensive, because the high resolution simulations have to be performed for all the different ensemble members.

5   Summary

We evaluate in detail the quality of the analyses of low latitudinal free tropospheric specific humidity (q), temperature (T), and

400   water vapour isotopologue ratio (δD), as well as of the three water cycle variables free tropospheric latent heating rate (Q_2), free tropospheric vertical velocity (ω), and precipitation rate (Prcp). We investigate the impact of assimilating free tropospheric specific humidity and temperature (which can be easily observed by many different techniques) and the possibility of further improving the analyses by additionally assimilating free tropospheric water isotopologue data (δD), for which nowadays also reliable observations with good horizontal and temporal coverage exist.

405   First, we make a statistical overview evaluation considering all locations and time steps. The assimilation of q and T observations strongly improves the analyses data quality of q and T with skill values of up to 60%, if compared to no data assimilation. Concerning the analyses of the other variables (δD, ω, Q_2, and Prcp) we also achieve a strong improvement with skill values of 30%-40%. Assimilating δD on top of q and T does further improve the analyses of δD (very strongly, with additional skills of up to 25%), of q and Prcp (with additional skills of about 5% and 10%, respectively), and of T and ω (by less than 10%

410   additional skills and an estimated skill uncertainty being slightly larger than the estimated skill value).

In a second evaluation we investigate how the additional assimilation impact of δD depends on the occurrence of phase transitions (condensation and evaporation leaves unique signatures on δD). We use atmospheric latent heating (Q_2) as proxy for phase transitions. The analyses uncertainty of ω, Q_2, and Prcp is the highest when strong latent heating occurs. We demonstrate that under these conditions the additional assimilation of δD on top of q and T has also the largest additional skills (up to 15%

415   for ω, Q_2, and Prcp). Because these strong latent heating events are very rare, the strong impact of the additional δD data assimilation is less clear in an overview evaluation, that considers all events (the large majority of events corresponds to heating rates close to zero, where the addition of δD observations have a very weak impact). However, the rare but strong latent heating events dominate the total yearly-averaged precipitation amounts in many regions and they are also responsible for extreme events (e.g. storms, flooding). Better capturing these events in the analyses has thus a strong societal impact.

420   For high heating rates, the additional assimilation of δD observations can also further improve the analyses of free tropospheric q and T. This improvement is high in relative metrics (skill values), but small in absolute values, because already the assimilation of q and T well constrains the q and T analyses.

The conditions when the additional assimilation of δD observations makes an important and unique contribution can be identified in $\{q, \delta D\}$-pairs located in the super-Rayleigh domain. These super-Rayleigh $\{q, \delta D\}$-pair distributions are often
425   observed in the MUSICA IASI satellite data. We interpret this as a promising indication for the possibility of improving the analyses in the real world similarly to the improvement demonstrated here with our OSSE for a simulated world. Moreover, the availability of a growing number of high resolution atmospheric isotopologue enabled models does importantly support further progress in this field.

Author contributions. Kei Yoshimura developed the isotopologue assimilation framework. Kinya Toride made all the data assimilation experiments. Matthias Schneider developed the ideas for evaluating the analyses improvements achievable by adding δD observations and made the respective calculations, whereby he was supported by Kinya Toride and Farahnaz Khosrawi. Frank Hase, Benjamin Ertl, and Christo-
435   pher J. Diekmann provided important contribution for the design of this study. All authors supported the generation of the final version of this manuscript.

Competing interests. At least one of the (co-)authors is a member of the editorial board of Atmospheric Measurement Techniques.

References

445 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, Journal of Geophysical Research: Atmospheres, 120, 9119–9137, https://doi.org/https://doi.org/10.1002/2015JD023403, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JD023403, 2015.

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, Journal of Geophysical Research: Atmospheres, 115,

450 https://doi.org/https://doi.org/10.1029/2010JD014554, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010JD014554, 2010.

Boesch, H., Deutscher, N. M., Warneke, T., Byckling, K., Cogan, A. J., Griffith, D. W. T., Notholt, J., Parker, R. J., and Wang, Z.: HDO/H$_2$O ratio retrievals from GOSAT, Atmospheric Measurement Techniques, 6, 599–612, https://doi.org/10.5194/amt-6-599-2013, http://www.atmos-meas-tech.net/6/599/2013/, 2013.

Bony, S., Stevens, B., Frierson, D., Jakob, C., Kageyama, M., Pincus, R., Shepherd, T. G., Sherwood, S. C., Siebesma, A. P.,

455 Sobel, A. H., Watanabe, M., and Webb, M. J.: Clouds, circulation and climate sensitivity, Nature Geosci, 8, 261–268, https://doi.org/https://doi.org/10.1038/ngeo2398, 2015.

Chan, S. C. and Nigam, S.: Residual Diagnosis of Diabatic Heating from ERA-40 and NCEP Reanalyses: Intercomparisons with TRMM, Journal of Climate, 22, 414 – 428, https://doi.org/https://doi.org/10.1175/2008JCLI2417.1, https://journals.ametsoc.org/view/journals/clim/22/2/2008jcli2417.1.xml, 2009.

460 Clerbaux, C., Boynard, A., Clarisse, L., George, M., Hadji-Lazaro, J., Herbin, H., Hurtmans, D., Pommier, M., Razavi, A., Turquety, S., Wespes, C., and Coheur, P.-F.: Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder, Atmospheric Chemistry and Physics, 9, 6041–6054, https://doi.org/10.5194/acp-9-6041-2009, https://acp.copernicus.org/articles/9/6041/2009/, 2009.

Craig, H.: Standard for Reporting concentrations of Deuterium and Oxygen-18 in natural waters, Science, 13, 1833–1834, https://doi.org/10.1126/science.133.3467.1833, 1961.

465 Dahinden, F., Aemisegger, F., Wernli, H., Schneider, M., a. D. 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, Atmospheric Chemistry and Physic, 21, 16 319–16 347, https://doi.org/10.5194/acp-21-16319-2021, 2021.

Diekmann, C. J.: Analysis of stable water isotopes in tropospheric moisture during the West African Monsoon, Ph.D. thesis, Karlsruher Institut für Technologie (KIT), https://doi.org/10.5445/IR/1000134744, 2021.

470 Diekmann, C. J., Schneider, M., Ertl, B., Hase, F., García, O., Khosrawi, F., Sepúlveda, E., Knippertz, P., and Braesicke, P.: The global and multi-annual MUSICA IASI {H$_2$O, δD} pair dataset, Earth System Science Data, 13, 5273–5292, https://doi.org/10.5194/essd-13-5273-2021, https://essd.copernicus.org/articles/13/5273/2021/, 2021a.

Diekmann, C. J., Schneider, M., Knippertz, P., de Vries, A. J., Pfahl, S., Aemisegger, F., Dahinden, F., Ertl, B., Khosrawi, F., Wernli, H., and Braesicke, P.: A Lagrangian Perspective on Stable Water Isotopes During the West African Monsoon, Journal of Geophysical Research:

475 Atmospheres, 126, e2021JD034 895, https://doi.org/https://doi.org/10.1029/2021JD034895, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021JD034895, e2021JD034895 2021JD034895, 2021b.

Eckstein, J., Ruhnke, R., Pfahl, S., Christner, E., Diekmann, C., Dyroff, C., Reinert, D., Rieger, D., Schneider, M., Schröter, J., Zahn, A., and Braesicke, P.: From climatological to small-scale applications: simulating water isotopologues with ICON-ART-Iso (version 2.3), Geoscientific Model Development, 11, 5113–5133, https://doi.org/10.5194/gmd-11-5113-2018, https://gmd.copernicus.org/articles/

480 11/5113/2018/, 2018.

Evans, C., Wood, K. M., Aberson, S. D., Archambault, H. M., Milrad, S. M., Bosart, L. F., Corbosiero, K. L., Davis, C. A., Pinto, J. R. D., Doyle, J., Fogarty, C., Galarneau, T. J., Grams, C. M., Griffin, K. S., Gyakum, J., Hart, R. E., Kitabatake, N., Lentink, H. S., McTaggart-Cowan, R., Perrie, W., Quinting, J. F. D., Reynolds, C. A., Riemer, M., Ritchie, E. A., Sun, Y., and Zhang, F.: The Extratropical Transition of Tropical Cyclones. Part I: Cyclone Evolution and Direct Impacts, Monthly Weather Review, 145, 4317 – 4344, https://doi.org/10.1175/MWR-D-17-0027.1, https://journals.ametsoc.org/view/journals/mwre/145/11/mwr-d-17-0027.1.xml, 2017.

Eyre, J. R., Bell, W., Cotton, J., English, S. J., Forsythe, M., Healy, S. B., and Pavelin, E. G.: Assimilation of satellite data in numerical weather prediction. Part II: Recent years, Quarterly Journal of the Royal Meteorological Society, 148, 521–556, https://doi.org/https://doi.org/10.1002/qj.4228, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.4228, 2022.

Field, R. D., Jones, D. B. A., and Brown, D. P.: Effects of postcondensation exchange on the isotopic composition of water in the atmosphere, J. Geophys. Res., 115, D24 305, https://doi.org/10.1029/2010JD014334, 2010.

Field, R. D., Kim, D., LeGrande, A. N., Worden, J., Kelley, M., and Schmidt, G. A.: Evaluating climate model performance in the tropics with retrievals of water isotopic composition from Aura TES, Geophysical Research Letters, 41, 6030–6036, https://doi.org/10.1002/2014GL060572, http://dx.doi.org/10.1002/2014GL060572, 2014GL060572, 2014.

Fink, A. H., Pohle, S., Pinto, J. G., and Knippertz, P.: Diagnosing the influence of diabatic processes on the explosive deepening of extratropical cyclones, Geophysical Research Letters, 39, https://doi.org/10.1029/2012GL051025, http://dx.doi.org/10.1029/2012GL051025, l07803, 2012.

Frankenberg, C., Yoshimura, K., Warneke, T., Aben, I., Butz, A., Deutscher, N., Griffith, D., Hase, F., Notholt, J., Schneider, M., Schrejver, H., and Röckmann, T.: Dynamic processes governing lower-tropospheric HDO/H$_2$O ratios as observed from space and ground, Science, 325, 1374–1377, https://doi.org/10.1126/science.1173791, 2009.

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, Reviews of Geophysics, 54, 809–865, https://doi.org/https://doi.org/10.1002/2015RG000512, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015RG000512, 2016.

González, Y., Schneider, M., Dyroff, C., Rodríguez, S., Christner, E., García, O. E., Cuevas, E., Bustos, J. J., Ramos, R., Guirado-Fuentes, C., Barthlott, S., Wiegele, A., and Sepúlveda, E.: Detecting moisture transport pathways to the subtropical North Atlantic free troposphere using paired H$_2$O-δD in situ measurements, Atmospheric Chemistry and Physics, 16, 4251–4269, https://doi.org/10.5194/acp-16-4251-2016, https://acp.copernicus.org/articles/16/4251/2016/, 2016.

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, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/https://doi.org/10.1002/qj.3803, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3803, 2020.

Hunt, B. R., Kostelich, E. J., and Szunyogh, I.: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter, Physica D: Nonlinear Phenomena, 230, 112–126, https://doi.org/https://doi.org/10.1016/j.physd.2006.11.008, https://www.sciencedirect.com/science/article/pii/S0167278906004647, data Assimilation, 2007.

Kalman, R. E.: A New Approach to Linear Filtering and Prediction Problems, Journal of Basic Engineering, 82, 35–45, https://doi.org/10.1115/1.3662552, https://doi.org/10.1115/1.3662552, 1960.

Lacour, J.-L., Risi, C., Clarisse, L., Bony, S., Hurtmans, D., Clerbaux, C., and Coheur, P.-F.: Mid-tropospheric dD observations from
    IASI/MetOp at high spatial and temporal resolution, Atmospheric Chemistry and Physics, 12, 10 817–10 832, https://doi.org/10.5194/acp-
520   12-10817-2012, http://www.atmos-chem-phys.net/12/10817/2012/, 2012.

Lacour, J.-L., Flamant, C., Risi, C., Clerbaux, C., and Coheur, P.-F.: Importance of the Saharan heat low in controlling the North At-
    lantic free tropospheric humidity budget deduced from IASI δD observations, Atmospheric Chemistry and Physics, 17, 9645–9663,
    https://doi.org/10.5194/acp-17-9645-2017, https://acp.copernicus.org/articles/17/9645/2017/, 2017.

Ling, J. and Zhang, C.: Diabatic Heating Profiles in Recent Global Reanalyses, Journal of Climate, 26, 3307 – 3325,
525   https://doi.org/https://doi.org/10.1175/JCLI-D-12-00384.1, https://journals.ametsoc.org/view/journals/clim/26/10/jcli-d-12-00384.1.xml,
    2013.

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.,
530   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, Journal of Geophysical Re-
    search: Atmospheres, 116, D22 113, https://doi.org/https://doi.org/10.1029/2011JD015773, https://agupubs.onlinelibrary.wiley.com/doi/
    abs/10.1029/2011JD015773, 2011.

Pante, G. and Knippertz, P.: Resolving Sahelian thunderstorms improves mid-latitude weather forecasts, Nature Communications, 10, 3487,
535   https://doi.org/10.1038/s41467-019-11081-4, https://doi.org/10.1038/s41467-019-11081-4, 2019.

Pfahl, S., Wernli, H., and Yoshimura, K.: The isotopic composition of precipitation from a winter storm – a case study with the
    limited-area model COSMO$_{iso}$, Atmospheric Chemistry and Physics, 12, 1629–1648, https://doi.org/10.5194/acp-12-1629-2012, https:
    //acp.copernicus.org/articles/12/1629/2012/, 2012.

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

Risi, C., Noone, D., Worden, J., Frankenberg, C., Stiller, G., Kiefer, M., Funke, B., Walker, K., Bernath, P., Schneider, M., Bony, S., Lee,
    J., Brown, D., and Sturm, C.: Process-evaluation of tropospheric humidity simulated by general circulation models using water vapor
    isotopic observations. Part 2: an isotopic diagnostic to understand the mid and upper tropospheric moist bias in the tropics and subtropics,
545   J. Geophys. Res., 117, https://doi.org/10.1029/2011JD016623, 2012.

Rodgers, C.: Inverse Methods for Atmospheric Sounding: Theory and Praxis, World Scientific Publishing Co., Singapore, 2000.

Schneider, A., Borsdorff, T., aan de Brugh, J., Aemisegger, F., Feist, D. G., Kivi, R., Hase, F., Schneider, M., and Landgraf, J.: First data set
    of H$_2$O/HDO columns from the Tropospheric Monitoring Instrument (TROPOMI), Atmospheric Measurement Techniques, 13, 85–100,
    https://doi.org/10.5194/amt-13-85-2020, https://amt.copernicus.org/articles/13/85/2020/, 2020.

550 Schneider, M. and Hase, F.: Optimal estimation of tropospheric H$_2$O and δD with IASI/METOP, Atmos. Chem. Phys., 11, 11 207–11 220,
    https://doi.org/10.5194/acp-11-11207-2011, 2011.

Schneider, M., Yoshimura, K., Hase, F., and Blumenstock, T.: The ground-based FTIR network's potential for investigating the atmospheric
    water cycle, Atmos. Chem. Phys., 10, 3427–3442, http://www.atmos-chem-phys.net/6/3427/2010/, 2010.

Schneider, M., Wiegele, A., Barthlott, S., González, Y., Christner, E., Dyroff, C., García, O. E., Hase, F., Blumenstock, T., Sepúlveda, E.,
555   Mengistu Tsidu, G., Takele Kenea, S., Rodríguez, S., and Andrey, J.: Accomplishments of the MUSICA project to provide accurate,

long-term, global and high-resolution observations of tropospheric {$H_2O,\delta D$} pairs – a review, Atmospheric Measurement Techniques, 9, 2845–2875, https://doi.org/10.5194/amt-9-2845-2016, http://www.atmos-meas-tech.net/9/2845/2016/, 2016.

Schneider, M., Borger, C., Wiegele, A., Hase, F., García, O. E., Sepúlveda, E., and Werner, M.: MUSICA MetOp/IASI {$H_2O,\delta D$} pair retrieval simulations for validating tropospheric moisture pathways in atmospheric models, Atmospheric Measurement Techniques, 10,
560   507–525, https://doi.org/10.5194/amt-10-507-2017, 2017.

Schneider, M., Ertl, B., Diekmann, C. J., Khosrawi, F., Weber, A., Hase, F., Höpfner, M., García, O. E., Sepúlveda, E., and Kinnison, D.: Design and description of the MUSICA IASI full retrieval product, Earth System Science Data, 14, 709–742, https://doi.org/10.5194/essd-14-709-2022, https://essd.copernicus.org/articles/14/709/2022/, 2022.

Sherwood, S. C., Bony, S., and Dufresne, J.-L.: Spread in model climate sensitivity traced to atmospheric convective mixing, Nature, 505,
565   37–42, https://doi.org/10.1038/nature12829, 2014.

Tada, M., Yoshimura, K., and Toride, K.: Improving weather forecasting by assimilation of water vapor isotopes, Scientific Reports, 11, https://doi.org/10.1038/s41598-021-97476-0, 2021.

Toride, K., Yoshimura, K., Tada, M., Diekmann, C., Ertl, B., Khosrawi, F., and Schneider, M.: Potential of Mid-tropospheric Water Vapor Isotopes to Improve Large-Scale Circulation and Weather Predictability, Geophysical Research Letters, 48, e2020GL091698,
570   https://doi.org/10.1029/2020GL091698, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020GL091698, 2021.

Webster, C. R. and Heymsfield, A. J.: Water Isotope Ratios H/D, $^{18}O/^{16}O$, $^{17}O/^{16}O$ in and out of Clouds Map Dehydration Pathways, Science, 302, 1742–1745, https://doi.org/10.1126/science.1089496, 2003.

Werner, M., Langebroek, P. M., Carlsen, T., Herold, M., and Lohmann, G.: Stable water isotopes in the ECHAM5 general circulation model: Toward high-resolution isotope modeling on a global scale, Journal of Geophysical Research: Atmospheres, 116,
575   https://doi.org/https://doi.org/10.1029/2011JD015681, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JD015681, 2011.

Wilks, D. S.: Statistical methods in the atmospheric sciences, Elsevier, https://doi.org/10.1016/C2017-0-03921-6, 2019.

Worden, J., Noone, D., Bowman, K., Beer, R., Eldering, A., Fisher, B., Gunson, M., Goldman, A., Herman, R., Kulawik, S. S., Lampel, M., Osterman, G., Rinsland, C., Rodgers, C., Sander, S., Shephard, M., Webster, R., and Worden, H.: Importance of rain evaporation and continental convection in the tropical water cycle, Nature, 445, 528–532, https://doi.org/10.1038/nature 05508, 2007.

580   Worden, J. R., Kulawik, S. S., Fu, D., Payne, V. H., Lipton, A. E., Polonsky, I., He, Y., Cady-Pereira, K., Moncet, J.-L., Herman, R. L., Irion, F. W., and Bowman, K. W.: Characterization and evaluation of AIRS-based estimates of the deuterium content of water vapor, Atmospheric Measurement Techniques, 12, 2331–2339, https://doi.org/10.5194/amt-12-2331-2019, https://amt.copernicus.org/articles/12/2331/2019/, 2019.

Yanai, M., Esbensen, S., and Chu, J.-H.: Determination of Bulk Properties of Tropical Cloud Clusters from Large-Scale Heat and Moisture
585   Budgets, Journal of Atmospheric Sciences, 30, 611 – 627, https://doi.org/10.1175/1520-0469(1973)030<0611:DOBPOT>2.0.CO;2, https://journals.ametsoc.org/view/journals/atsc/30/4/1520-0469_1973_030_0611_dobpot_2_0_co_2.xml, 1973.

Yoshimura, K., Kanamitsu, M., Noone, D., and Oki, T.: Historical isotope simulation using Reanalysis atmospheric data, J. Geophys. Res., 113, D19 108, https://doi.org/10.1029/2008JD010074, 2008.

Yoshimura, K., Miyoshi, T., and Kanamitsu, M.: Observation system simulation experiments using water vapor isotope information,
590   Journal of Geophysical Research: Atmospheres, 119, 7842–7862, https://doi.org/https://doi.org/10.1002/2014JD021662, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014JD021662, 2014.