Atmos. Chem. Phys., 22, 15559–15577, 2022
https://doi.org/10.5194/acp-22-15559-2022

Atmospheric Chemistry and Physics
Open Access
EGU

Research article

Vertical structure of the lower-stratospheric moist bias in the ERA5 reanalysis and its connection to mixing processes

Konstantin Krüger[1], Andreas Schäfler[1], Martin Wirth[1], Martin Weissmann[3], and George C. Craig[2]

[1]Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany
[2]Meteorologisches Institut München, Ludwig-Maximilians-Universität, Munich, Germany
[3]Institut für Meteorologie und Geophysik, Universität Wien, Vienna, Austria

Correspondence: Konstantin Krüger (konstantin.krueger@dlr.de)

Received: 15 July 2022 – Discussion started: 22 July 2022
Revised: 11 November 2022 – Accepted: 17 November 2022 – Published: 12 December 2022

Abstract. Numerical weather prediction (NWP) models are known to possess a distinct moist bias in the mid-latitude lower stratosphere, which is expected to affect the ability to accurately predict weather and climate. This paper investigates the vertical structure of the moist bias in the European Centre for Medium-Range Weather Forecasts (ECMWF) latest global reanalysis ERA5 using a unique multi-campaign data set of highly resolved water vapour profiles observed with a differential absorption lidar (DIAL) on board the High Altitude and LOng range research aircraft (HALO). In total, 41 flights in the mid-latitudes from six field campaigns provide roughly 33 000 profiles with humidity varying by 4 orders of magnitude. The observations cover different synoptic situations and seasons and thus are suitable to characterize the strong vertical gradients of moisture in the upper troposphere and lower stratosphere (UTLS). The comparison to ERA5 indicates high positive and negative deviations in the UT, which on average lead to a slightly positive bias (15 %–20 %). In the LS, the moist bias rapidly increases up to a maximum of 55 % at 1.3 km altitude above the thermal tropopause (tTP) and decreases again to 15 %–20 % at 4 km altitude. Such a vertical structure is frequently observed, although the magnitude varies from flight to flight. The layer depth of increased moist bias is smaller at high tropopause altitudes and larger when the tropopause is low. Our results also suggest a seasonality of the moist bias, with the maximum in summer exceeding autumn by up to a factor of 3. During one field campaign, collocated ozone and water vapour profile observations enable a classification of tropospheric, stratospheric, and mixed air using water vapour–ozone correlations. It is revealed that the moist bias is high in the mixed air while being small in tropospheric and stratospheric air, which highlights that excessive transport of moisture into the LS plays a decisive role for the formation of the moist bias. Our results suggest that a better representation of mixing processes in NWP models could lead to a reduced LS moist bias that, in turn, may lead to more accurate weather and climate forecasts. The lower-stratospheric moist bias should be borne in mind for climatological studies using reanalysis data.

Published by Copernicus Publications on behalf of the European Geosciences Union.

1 Introduction

Water vapour is one of the most important greenhouse gases in the atmosphere and plays a key role for accurately predicting the Earth's weather (Gray et al., 2014; Shepherd et al., 2018) and climate (Forster and Shine, 2002; Riese et al., 2012). In the upper troposphere and lower stratosphere (UTLS), defined as a layer located ±5 km around the thermal tropopause (tTP) (Gettelman et al., 2011), rapidly decreasing water vapour concentrations in the vertical (e.g. Kiemle et al., 2012; Kaufmann et al., 2018) are of key relevance to a net cooling near and above the tropopause (Randel et al., 2007). The radiative modulation of the vertical temperature gradients may influence the near-tropopause potential vorticity (PV) gradient (Chagnon et al., 2013) that acts as a waveguide for Rossby waves (Martius et al., 2010) and thus may affect downstream weather development in the mid-latitudes. Hence, an accurate representation of UTLS water vapour in numerical weather prediction (NWP) and climate models is essential.

In the extratropical UTLS, the distribution of water vapour is driven by transport and mixing processes related to baroclinic waves and associated synoptic- and meso-scale weather systems, which are interacting with chemical processes (e.g. Gettelman et al., 2011; Schäfler et al., 2022). The increased static stability above the tropopause (Birner et al., 2002) impedes water vapour from being vertically transported. Correspondingly, the sharpest decline of water vapour is found just above the tropopause. Exchange processes affect the water concentration around the tropopause (Holton et al., 1995; Stohl et al., 2003) and create the extratropical transition layer (ExTL; Pan et al., 2004; Hoor et al., 2010) with influences of the troposphere and the stratosphere. In particular quasi-isentropic exchange near the polar and subtropical jet streams (Haynes and Shuckburgh, 2000) and cross-isentropic mixing, for instance, through overshooting convection (e.g. Dessler and Sherwood, 2004; Homeyer et al., 2014), are major contributors to increased humidity above the tropopause. Furthermore, tropopause folds are related to mass exchange between the UT and the LS (Shapiro, 1980). Above the ExTL, the concentration of water vapour approaches a low and vertically constant background value (e.g. Hintsa et al., 1994), which is determined by the stratospheric transport from tropics (Fueglistaler et al., 2009) within the Brewer–Dobson circulation (e.g. Dobson et al., 1946; Brewer, 1949) on timescales from months to years (Birner and Bönisch, 2011). The complexity of transport and mixing processes is mirrored in the high water vapour variability in the extratropical UTLS on synoptic and seasonal timescales (e.g. Pan et al., 2000; Randel and Wu, 2010; Zahn et al., 2014; Dyroff et al., 2015; Bland et al., 2021; Schäfler et al., 2022).

The sharp vertical gradients of trace species, PV, wind, and temperature at the extratropical tropopause are challenging to resolve for state-of-the-art NWP models (e.g. Stenke et al., 2008; Schäfler et al., 2020). Current NWP analyses and forecasts are known to possess a distinct moist bias in the extratropical LS (e.g. Kaufmann et al., 2018), which is causing a collocated cold bias at the same altitudes (Stenke et al., 2008; Diamantakis and Flemming, 2014; Shepherd et al., 2018). Recently, Bland et al. (2021) used radiosonde observations of a 2-month period in autumn and confirmed the earlier documented moist bias (about 70 % in the LS) in current operational analyses of the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecast System (IFS) and the Met Office's Unified Model (METUM). They also showed that radiative effects related to the moist bias cause a collocated cold bias in the LS that grows with forecast lead time. For a comprehensive overview of the studies that quantified the LS moist bias in different NWP systems, the interested reader is referred to Table 1 in Bland et al. (2021). The vertical structure of the moist bias is characterized by a small positive bias below the thermal tropopause, followed by a vertical increase in the LS to a maximum at 1–2 km above the tropopause (e.g. Dyroff et al., 2015; Bland et al., 2021). However, different shapes of the LS moist bias above its maximum have been reported. Bland et al. (2021) show an opposing vertical structure of the moist bias beyond 2 km above the tTP using two different radiosonde types. Woiwode et al. (2020) compare humidity cross sections of an airborne passive infrared imager and present a case with vertically increasing moist bias, one with constant bias, and two cases with vertically decreasing moist bias in IFS analysis and forecast data.

The origin of the wet model bias is still under debate: one hypothesis is that it is caused by misrepresented dynamical transport and mixing processes (Kunz et al., 2014; Shepherd et al., 2018), for example, overshooting convection leading to excessive water vapour injection into the LS. Another potential source of overestimated transport of moisture into the LS is numerical diffusion and insufficient model resolution in the semi-Lagrangian advection scheme used in the ECMWF model, leading to an excessive diffusive transport of moisture across strong gradients from high to low mixing ratios (Stenke et al., 2008; Kunz et al., 2014; Dyroff et al., 2015; Shepherd et al., 2018). However, a LS moist bias of similar order is also found for "Eulerian"-formulated models (Jiang et al., 2015; Davis et al., 2017). Moreover, Woiwode et al. (2020) confirm that the bias is already present in the initial conditions and demonstrate a low response of the moist bias to variable vertical or temporal resolutions.

The above-mentioned studies used a variety of observation techniques to quantify the moist bias. Radiosonde or dropsonde humidity observations provide temporally continuous series of profiles at the same location, but their reliability is limited > 2 km above the tTP (e.g. Bland et al., 2021). In situ aircraft observations, even though very accurate and highly resolved, provide profile information only during start and landing and on flight routes of commercial or research aircraft (Zahn et al., 2014; Kunz et al., 2014; Dyroff et al.,

2015; Kaufmann et al., 2018). Conversely, spaceborne microwave sounders provide humidity information across the entire globe but are limited in vertical resolution (e.g. Hegglin et al., 2013; Jiang et al., 2015; Khosrawi et al., 2018). In between the in situ and satellite observations, profile data from active and passive remote sensing instruments on board research aircraft demonstrated the potential to characterize humidity across the tropopause (Ehret et al., 1999; Flentje et al., 2007; Woiwode et al., 2020; Schäfler et al., 2021), combining high spatial coverage, high accuracy, and high vertical resolution (Bhawar et al., 2011). Since 2013, the active Differential Absorption Lidar (DIAL) WAter vapour and Lidar Experiment in Space (WALES; Wirth et al., 2009) has been deployed in several research campaigns on board the High Altitude and LOng range research aircraft (HALO; Krautstrunk and Giez, 2012) for water vapour profile measurements.

The goal of this paper is to evaluate the LS moist bias in the ECMWF's most recent global reanalysis ERA5. The model analyses are compared against a comprehensive data set of water vapour profiles observed by the airborne DIAL WALES in the mid-latitude UTLS. Collocated water vapour and ozone profiles are used to identify tropospheric, stratospheric, and mixed air and to individually assess the moist bias as we suspect that mixing processes affect the vertical structure of the moist bias. The following three specific questions are addressed:

1. Can the multi-campaign DIAL data set robustly quantify the LS moisture bias in ERA5?
2. What is the vertical structure of the LS moist bias in ERA5, particularly at high altitudes?
3. Is the moist bias correlated to the distribution of mixed air masses in the UTLS?

This paper is outlined as follows: Sect. 2 provides an overview of the water vapour DIAL observations (Sect. 2.1), the ERA5 reanalysis (Sect. 2.2), and the methods utilized to compare the observational and model data (Sect. 2.3). In Sect. 3.1 an example cross section of specific humidity and the bias are illustrated for a mid-latitude jet stream crossing, which is followed by a statistical tropopause-relative evaluation of the vertical structure of the bias and its variability in Sect. 3.2. The relationship between the vertical structure of the moist bias and the distribution tropospheric, stratospheric, and mixed air is presented in Sect. 3. Thereafter, Sect. 4 provides a discussion of the results. The key conclusions are summarized in Sect. 5.

2 Data and methods

2.1 The WALES data set

The DIAL WALES (Wirth et al., 2009) was developed at the German Aerospace Center (DLR) and has been operated on board the German research aircraft HALO since 2010. The instrument design is based on two identical laser systems that generate four wavelengths in the near-infrared (NIR) absorption band of water vapour between 935 and 936 nm, allowing for water vapour observations from the planetary boundary layer up to the stratosphere. WALES furthermore operates two polarization-sensitive channels at 1064 nm and at 532 nm. The latter channel comprises a high spectral resolution lidar (HSRL; Esselborn et al., 2008), enabling extinction coefficient observations and thus aerosol characterization (Groß et al., 2013). WALES and its underlying DIAL technique is briefly introduced in the following, and a more detailed description can be found in Wirth et al. (2009).

The four NIR wavelengths are separated into three online channels (strongly absorbed by water vapour) and one offline channel (weakly absorbed). The number concentration of water vapour in the probed volume is derived from the ratio of the backscattered light of the on- and offline wavelengths and then converted to specific humidity. The online channels are sensitive to different trace gas concentrations and in turn to different altitude levels. The exact wavelengths are selected such that they are optimally aligned to the moist boundary layer, the UT, and the dry LS. Note that the WALES humidity profiles are only available in cloud-free regions or regions with optically thin clouds. In optically thick clouds the extinction by cloud particles is so strong that no water vapour information can be retrieved within or below the cloud.

Due to the photon statistics of the backscattered light as well as detector and background light noise, the retrieved water vapour profiles undergo statistical variations, which are effectively reduced by temporal (i.e. horizontal) and vertical averaging. Thus, the retrieved DIAL water vapour profiles are averaged over 12 s or approximately 3 km in the horizontal. In the vertical, data are available every 15 m, although the effective vertical resolution is 300 m according to the full width of half maximum of the averaging kernel. It should be stressed that the averaging kernel of the WALES DIAL is exactly zero outside of about $\sqrt{2}$ times the effective resolution. This is in sharp contrast to most passive remote sensing techniques, where the side modes of the kernels can lead to erroneous dry or wet layers in the retrieved humidity profile. In the DIAL data retrieval, the statistical error of the observed volume is different for each flight and depends on the water vapour distribution and the background light. To remove high noise, typically occurring in dry air lying underneath moist air, for example, in the vicinity of stratospheric intrusions (Trickl et al., 2016), we filtered 5 % of the noisiest data for each individual flight. This threshold turned out to be useful but reduced the data availability in the lower- to mid-troposphere. Furthermore, Rayleigh–Doppler beam broadening, laser spectral impurity, and uncertainties in spectral databases are sources of systematic errors, which are compensated for in the retrieval algorithm. The total systematic error was found to be in the order of 5 % (Kiemle et

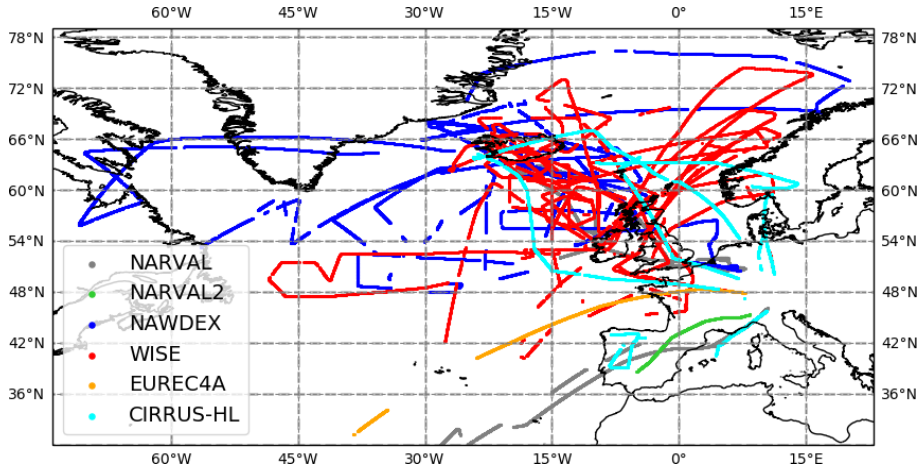

Figure 1. Map of HALO flight sections with WALES DIAL water vapour observations during the research campaigns NARVAL, NARVAL2, NAWDEX, WISE, EUREC^{4}A, and CIRRUS-HL (for a detailed overview, see Sect. 2.1).

Table 1. Overview of all considered campaigns with DIAL observations. The number of DIAL profiles refers to all profiles that were sampled during 41 flights. The number of DIAL profiles in the LS corresponds to all profiles with measurements in the LS.

Campaign	Month and year	Season	No. of flights (hours)	Flight distance (km)	No. of DIAL profiles	No. of DIAL profiles in the LS
NARVAL	Dec–Jan 2013/2014	Winter	7 (41)	31 157	10 973	5288
NARVAL2	Aug 2016	Summer	1 (9)	7729	2395	485
NAWDEX	Sep–Oct 2016	Autumn	11 (75)	55 695	19 139	12 062
WISE	Sep–Oct 2017	Autumn	14 (105)	83 041	13 557	9493
EUREC^{4}A	Jan–Feb 2020	Winter	1 (8)	7011	2307	1009
CIRRUS-HL	Jun–Jul 2021	Summer	7 (30)	23 675	6777	4568
Total			41 (268)	208 308	55 148	32 905

al., 2008). The high reliability of WALES was demonstrated in various intercomparisons, for example, with Lyman-α in situ hygrometers (Kiemle et al., 2008), comparable airborne and ground-based DIAL instruments (Bhawar et al., 2011), and radiosondes with a frost point hygrometer (Trickl et al., 2016).

In this study, we use DIAL observations from six campaigns from 2013–2021 that provide almost 33 000 water vapour profiles obtained during 41 research flights. The profiles were sampled along the flight track and extend from the surface up to about 14 km altitude, corresponding to the maximum flight level of the HALO aircraft (Krautstrunk and Giez, 2012). As the focus of this study is the mid-latitude UTLS, we only consider flights that provide a significant amount of data across the tropopause. The majority (25) of these flights took place in the northern hemispheric autumn season during the North Atlantic Waveguide Downstream impact EXperiment (NAWDEX; Schäfler et al., 2018) and the Wave-driven ISentropic Exchange campaign (WISE; Kunkel et al., 2019). As part of the campaigns ElUcidating the RolE of Cloud-Circulation Coupling in ClimAte (EUREC^{4}A; Stevens et al., 2021), the Next-generation Aircraft Remote sensing for VALidation studies (NARVAL; Klepp et al., 2014), and NARVAL2 (Stevens et al., 2019) measurements were taken during eight flights in the winter season. In addition, the Cirrus in High-Latitudes (CIRRUS-HL) mission provides observations in summer. Figure 1 depicts the parts of HALO research flights where DIAL observations were obtained. Most flights were carried out over the North Atlantic between 48 and 66° N, the North Sea and central to western Europe. Additionally, the subtropics (> 35° N) and the Arctic were covered by individual flights as well.

During the WISE campaign, WALES was operated in a different setup to measure both water vapour and ozone, concurrently. For this purpose, two of the 935 nm NIR water vapour channels were replaced by two ultraviolet (UV) channels covering the 300–305 nm ozone absorption line (Fix et al., 2019). The use of two instead of four channels per trace gas leads to a reduced vertical coverage, which was optimized so that the selected NIR wavelengths cover the tropopause region. Increased statistical noise required averaging over a period of 24 s (∼ 6 km horizontally), while

the effective vertical resolution remains approximately 300 m (Fix et al., 2019).

The number of observations with respect to latitude (Fig. 2a) illustrates the high data availability in the mid-latitudes, which is the region of interest in this study. This data set that covers humidity observations in a broad spectrum of synoptic situations is considered to be representative of mid-latitude weather. Figure 2b shows the distribution of the water vapour observations covering 4 orders of magnitude, ranging from 10^{-3} to 10^{1} g kg^{-1}. The bimodal shape of the histogram is composed of a broad moist part that can be assigned to the troposphere and a fraction of low humidity representing the dry conditions in the stratosphere. Each campaign exhibits an individual footprint of measured humidity, depending on the season, observation areas, and the flight level selection. For instance, the histograms for NAWDEX and WISE are remarkably similar since both campaigns took place over the North Atlantic in autumn. However, as only two NIR wavelengths were operated to measure water vapour during WISE, fewer measurements are available at high humidity levels. NARVAL shows a distinctive dry spectrum of measured humidity corresponding to the winter season, and fewer data are available for the LS, resulting from frequent low flight altitudes. The CIRRUS-HL summer campaign stands out because a large proportion of high moisture values was observed. The NARVAL2 and the EUREC^{4}A campaign provide UTLS measurements only for one flight and, thus, compared to the other field campaigns, provide a small number of observations (see also Table 1).

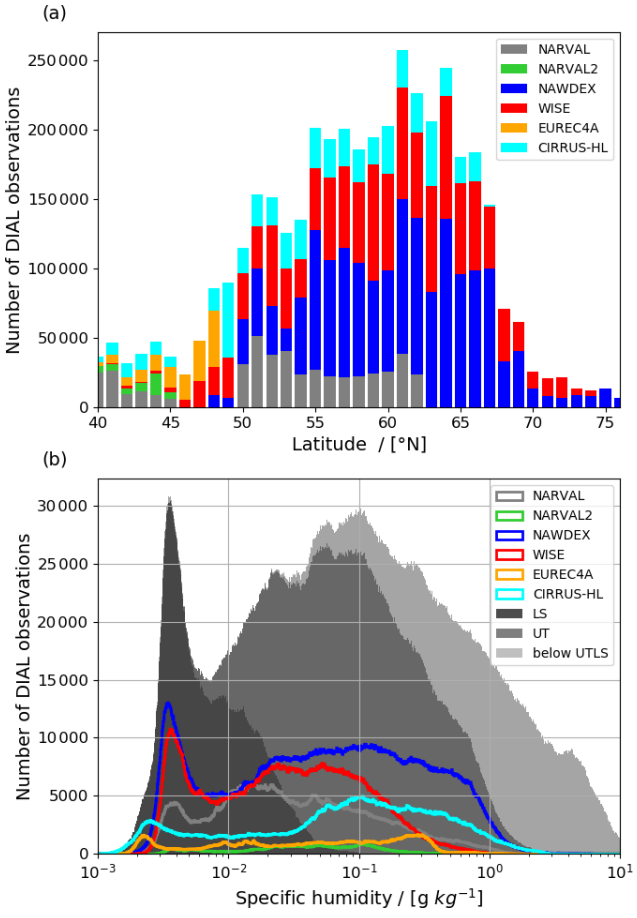

Figure 2. **(a)** Stacked distribution of the number of observations in 1° latitude bins per individual campaign (coloured bars). **(b)** Histogram of observations per humidity bin with a size of 0.01 g kg^{-1} of $\log_{10}$ (q_{DIAL}) for individual campaigns in the UTLS (coloured lines). Shading shows frequencies separated for the LS (all data above the thermal tropopause, dark grey shading), the UT (all data between the tTP and 5 km below, medium grey shading), and the remaining tropospheric data (light grey shading).

2.2 ERA5 reanalysis data

ERA5 is the latest-generation reanalysis of the ECMWF based on the IFS Cycle 41r2 that was used for operational weather prediction in 2016. Atmospheric quantities are provided on a global grid with a horizontal resolution (TL639) of about 31 km and on 137 hybrid sigma-pressure model levels, ranging from the surface up to 0.01 hPa (∼ 80 km) in the vertical. The altitude range of the DIAL observations is covered by the lowermost 70 model levels. The vertical grid spacing of the model levels ranges from a few metres in the boundary layer to about 300 m at the tropopause level (Schäfler et al., 2020). ERA5 reanalyses are available with a time resolution of 1 h, which is an improvement compared to a 6-hourly resolution of its predecessor ERA-Interim (Dee et al., 2011). Further details about ERA5 are documented in Hersbach et al. (2020). For this study, model level data are retrieved on a regular 0.36° × 0.36° latitude–longitude grid. Pressure and altitude of each model level are derived following the IFS documentation (ECMWF, 2015). To be able to compare ERA5 and WALES data, the gridded model data are interpolated in space and time to the observation location. Our interpolation method uses a horizontally bilinear interpolation, followed by a linear interpolation in the vertical. Finally, a linear interpolation in time of the hourly ERA5 profiles towards the observation time is carried out. This sequence of interpolation steps has been applied similarly in other studies (e.g. Schäfler et al., 2010).

2.3 Data processing

2.3.1 Thermal tropopause detection

Due to the variable altitude of the tTP, the distribution of water vapour in the UTLS at individual altitudes is also highly variable. Hence, averaging of the humidity profiles in geometrical coordinates strongly blurs the vertical gradients across the tropopause. Therefore, bias statistics are often performed in tropopause-relative coordinates (e.g. Kunz et al., 2014; Bland et al., 2021). Different tropopause definitions have been established, taking the thermal, dynamical, and chemical properties of the UTLS as a reference. By defini-

Table 2. Some example values of specific humidity and the according computed humidity bias.

Q_{ERA5}	$\mathrm{g\,kg^{-1}}$	0.50	0.75	1.00	1.25	1.50	1.75	2.00	2.25	2.50	3.00
Q_{DIAL}	$\mathrm{g\,kg^{-1}}$	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Humidity bias	Unitless	−1.00	−0.41	0	0.32	0.58	0.81	1.00	1.17	1.32	1.58
Percentage	%	−50	−25	0	25	50	75	100	125	150	200

tion, the tTP marks the reversal of the vertical temperature gradient and thus the abrupt increase in static stability, which is reflected in the sharp distribution of trace species across the tropopause (Gettelman et al., 2011). We use the tTP as it best reflects the strongest vertical gradients of water vapour (Birner et al., 2002; Pan et al., 2004). From each ERA5 temperature profile interpolated to the 15 m vertical grid of the lidar, we calculate the tTP altitude using the World Meteorological Organization's (WMO) lapse rate-based definition (WMO, 1957). A tTP is detected as the lowest level at which the vertical temperature gradient Γ drops below $2\,\mathrm{K\,km^{-1}}$ and is only defined if the average lapse rate between this and any other level within a 2 km deep layer remains equal or lower than $2\,\mathrm{K\,km^{-1}}$. The vertical temperature gradient, i.e. the lapse rate, is computed as

$$\Gamma = \left(-\frac{\mathrm{d}T}{\mathrm{d}z}\right)\,[\mathrm{K\,km^{-1}}]. \tag{1}$$

In our analyses, the tTP detection is started in an upward direction from 5 km altitude in order to avoid misdetections of tropopauses due to local fluctuations of temperature in the lower- to mid-troposphere. When a tTP is detected, the (thermal) tropopause-relative coordinates $z_{\mathrm{rel.tTP}}$ are derived by simply subtracting the altitude of tTP (z_{tTP}) from the geometric height vector (z_{geom}):

$$z_{\mathrm{rel.\,tTP}} = z_{\mathrm{geom}} - z_{\mathrm{tTP}}. \tag{2}$$

There are conditions in which tropopause detection is ambiguous, especially in the vicinity of the jet streams and associated tropopause folds, where double tropopauses can occur (e.g. Shapiro, 1980; Gettelman et al., 2011). We found that in situations of weak vertical temperature gradients near the jet streams, the lapse rate threshold in the WMO definition may lead to vertical jumps of the tTP altitudes for adjacent profiles. These fluctuations result in a wrong vertical allocation of water vapour in tropopause-relative coordinates. A detailed discussion will follow in Sect. 3.1. To remove such profiles in the overall statistic, we apply a filtering method based on mean potential vorticity (MPV; Shapiro et al., 1999), which is the average PV calculated for the 5 km layer above and below the thermal tropopause. MPV < 3.5 potential vorticity units (PVU; $1\,\mathrm{PVU} = 10^{-6}\,\mathrm{K\,m^2\,kg^{-1}\,s^{-1}}$) above and MPV > 3.5 PVU below the tTP are found to be an efficient metric to filter profiles within an erroneously assigned tTP. Figure 3 shows the vertical distribution of tTP altitudes for the 32 905 profiles, which lies between 5.5 km and more than

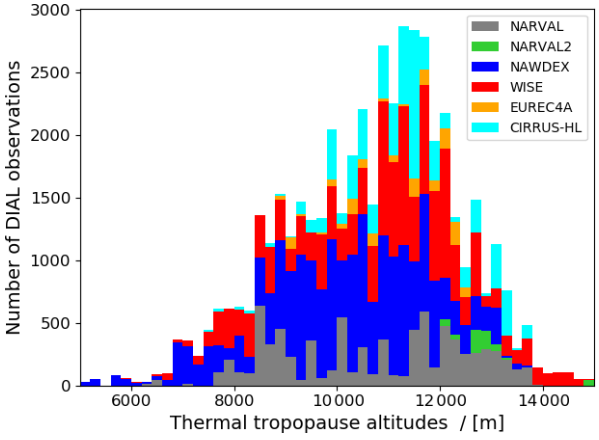

Figure 3. Histogram of the number of observations per thermal tropopause altitude bin (1000 m) and per campaign (coloured bars).

15 km altitude, reflecting the broad spectrum of synoptic situations covered by the data set. The majority of all tTP levels are found between 10 and 13 km, which represents the typical location of the mid-latitude tropopause with respect to interannual or synoptic variations (e.g. Birner et al., 2002).

2.3.2 Statistical metric of the bias

The selection of a suitable difference metric is crucial for a robust quantification of model humidity errors, and different statistical approaches can be found in the literature (Kunz et al., 2014; Bland et al., 2021). As specific humidity rapidly decreases across the tropopause, absolute humidity differences are not appropriate, and most studies rely on a relative formulation of the error. However, since the simple ratio of model and observation as well as the absolute bias divided by the observed value are statistical asymmetric quantities, we apply a logarithmic formulation with base 2 (see Eq. 3), introduced by Kunz et al. (2014):

$$\text{humidity bias} = \log_2\left(\frac{Q_{\mathrm{ERA5}}}{Q_{\mathrm{DIAL}}}\right), \tag{3}$$

with Q_{DIAL} being the measured and Q_{ERA5} being the ERA5 specific humidity. This unitless definition of the relative bias is symmetrically centred around zero and thus not distorted when averaged. A perfect agreement (humidity bias = 0) between the ERA5 and the DIAL specific humidity is reached if $Q_{\mathrm{ERA5}} = Q_{\mathrm{DIAL}}$. A positive humidity bias $\epsilon[0, \infty]$ indicates an overestimation of humidity by the model ($Q_{\mathrm{ERA5}} >$

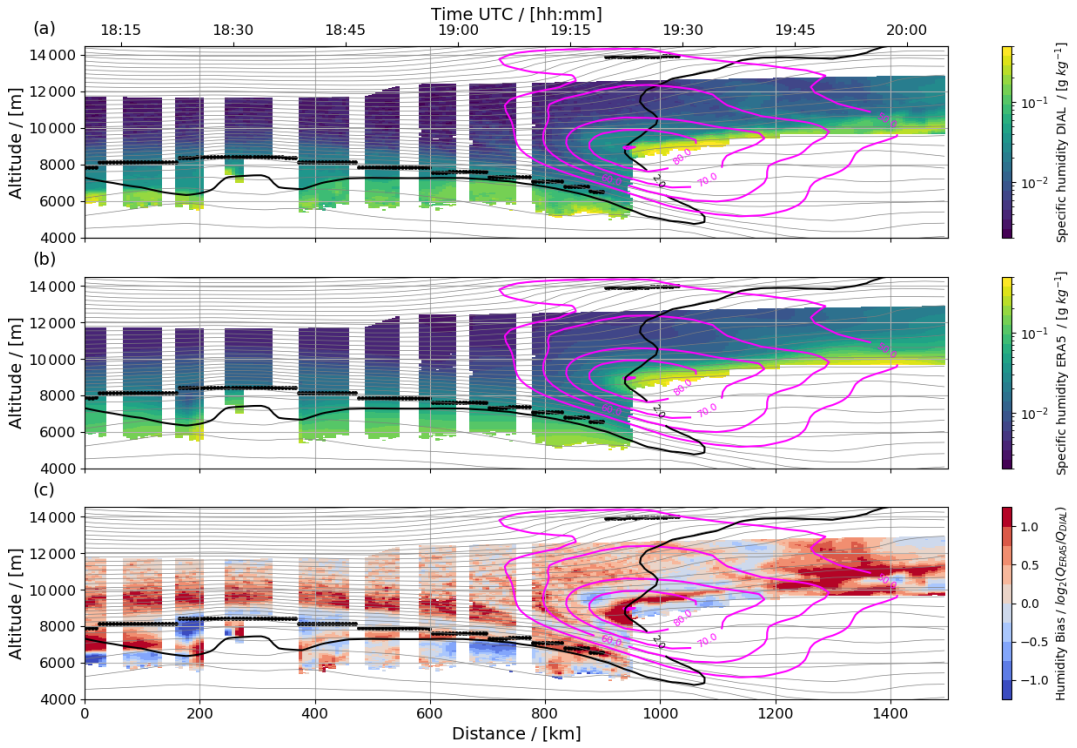

Figure 4. Vertical cross sections of **(a)** the DIAL specific humidity (colour shading, g kg^{-1}), **(b)** ERA5 specific humidity (colour shading, g kg^{-1}), and **(c)** the corresponding humidity bias (colour shading) on 1 October 2017. Panels **(a)**–**(c)** are superimposed by ERA5 fields of the potential temperature (grey contours, $\Delta\theta = 3$ K), the isopleths of the wind speed (magenta contours, in m s^{-1}), and the thermal (thick black dots) and the dynamical tropopause (2 PVU, black isoline).

Q_{DIAL}), whereas a negative humidity bias $\epsilon[-\infty, 0]$ implies an underestimation ($Q_{\mathrm{ERA5}} < Q_{\mathrm{DIAL}}$). Table 2 gives some example bias values for selected moisture observations.

3 Results

3.1 Water vapour and bias distributions for a selected case

First, an example cross section of water vapour measurements of the research flight on 1 October 2017 during the WISE campaign is presented in Fig. 4. The case is selected as it possesses a good data coverage across the UTLS and as it additionally provides ozone observations (see Sect. 3.3). HALO flew meridional transects over the North Atlantic (50–60° N) at 13° W, aiming to measure a zonal jet stream and its associated predicted strong trace gas gradients. The underlying synoptic situation and the corresponding mission objectives are provided in detail by Schäfler et al. (2021). The left part of Fig. 4a (up to a distance of roughly 800 km) illustrates the water vapour distribution north of the jet stream (see magenta isopleths) where the aircraft flew above the low-located tropopause within the LS. HALO then crossed the pronounced jet stream with wind velocities of more than 90 m s^{-1}. Near the jet core, the tTP altitude jumps from 6.5 to 14 km within a few kilometres' flight distance. The dynamical tropopause (2 PVU contour line) also displays the ascent of the tropopause and a corresponding tropopause fold that extends along inclined isentropes into the mid-troposphere. In the right part of Fig. 4a, the air mass located to the south of the jet stream exhibits high tropopause altitudes exceeding the flight level by roughly 2 km, so that measurements are restricted to tropospheric air. Along the entire cross section, the highest specific humidity is observed at the lowest levels in the UT, ranging from 10^{-2} g kg^{-1} to occasionally more than 10 g kg^{-1}. The tropospheric air to the south of the jet stream has an increased humidity content compared to the air north of the jet stream. In the LS, specific humidity values lower than 10^{-2} g kg^{-1} are frequently observed. At a first glance, the specific humidity curtain of ERA5 (Fig. 4b) is very similar to the observations. However, the ERA5 humidity field appears to be smoother, particularly in the presence of strong horizontal water vapour gradients, for instance, near the jet stream and mesoscale filaments. Differences between observations and model, calculated by applying Eq. (3), are shown for the vertical section in Fig. 4c. Reddish regions indicate an overestimation of humidity by ERA5, while bluish areas represent an underestimation. High positive and negative values of the bias alternate below the tropopause. In the LS, a coherent region of positive values is detected between 1 and 3 km

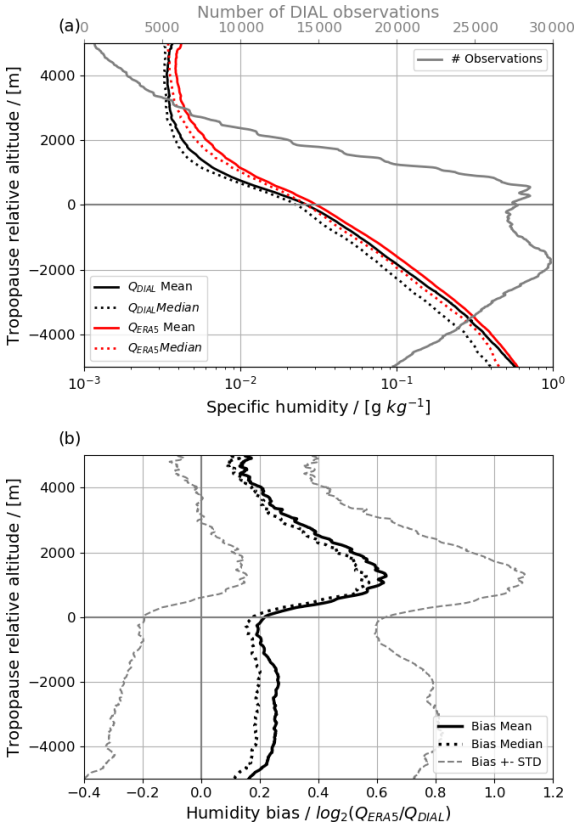

Figure 5. Tropopause-relative **(a)** vertical profiles of the DIAL (black lines) and the ERA5 (red lines) mean (solid) and median (dotted) specific humidity and the number of observations (grey). Note the log-scale notation of the x axis. **(b)** Mean/median bias (solid/dotted lines) and standard deviation (dotted grey lines).

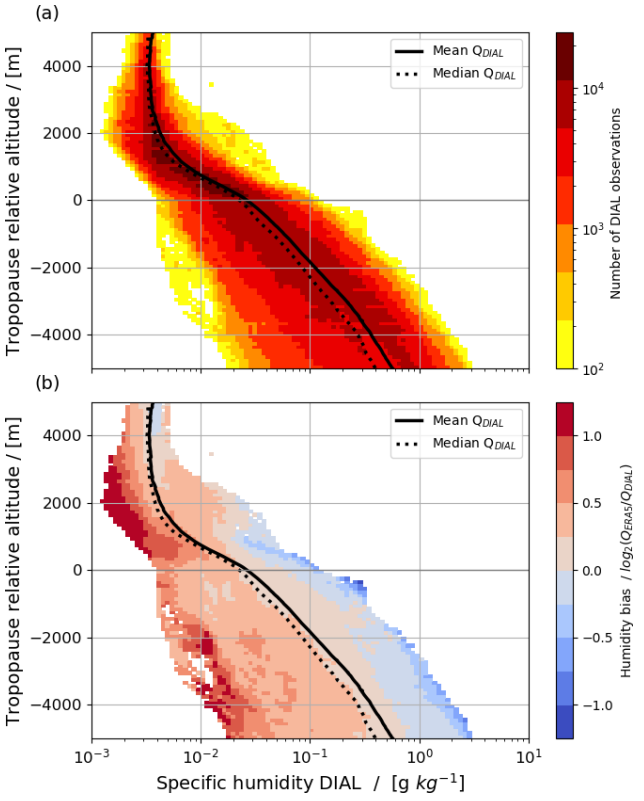

Figure 6. Binned distribution of DIAL specific humidity observations relative to the thermal tropopause coloured by **(a)** the number of observations per bin and **(b)** the bin-average humidity bias. Solid thick black (dotted) lines in **(a)** and **(b)** show mean (median) values per altitude bin. Bin sizes are 100 m and 0.01 g kg^{-1} of $\log_{10}$ (q_{DIAL}). Please note the logarithmic abscissa and colour bar in **(a)**.

above the tTP, indicating an overestimated humidity that extends over the entire part north of the jet. At the highest altitudes, beyond 3 km above the tropopause, the moist bias is smaller. In order to study the systematic nature of the diagnosed LS moist bias and its vertical structure, a statistic of all observations in tropopause-relative coordinates is performed.

3.2 Statistical analysis of the LS bias

3.2.1 Vertical structure

For all 32 905 profiles from the 41 flights, the average profiles of specific humidity and the humidity bias are presented in Fig. 5. The moisture profiles of WALES and ERA5 show an exponential decline of specific humidity in the UT, ranging from about 5×10^{-1} g kg^{-1} at the lowest levels to approx. 3×10^{-2} g kg^{-1} at the tropopause. The strongest vertical gradient occurs in a layer of 0.5 to 1 km above the tropopause. Beyond, a less pronounced decline of water vapour extends until 4 km above the tropopause, followed by a vertical constant specific humidity of about 3.5×10^{-3} g kg^{-1}. There is a high level of agreement between the ERA5 and WALES specific humidity profiles, particularly in the UT, although ERA5 appears to be moister at all altitudes. For both data sets, the median and arithmetic mean profiles of specific humidity slightly vary from each other. The median line is slightly shifted towards drier humidity values, most pronounced in the UT. Figure 5a also demonstrates a high data availability throughout the entire UTLS. The number of observations is highest between −5 and 1 km around the tTP, with two local maxima at −2 km and roughly 1 km. Note that these two peaks in data availability are related to the typical flight altitudes, either above or below the main transatlantic air traffic routes (Schäfler et al., 2018), and the maximum data coverage close to the aircraft. Above the tTP, the number of observations continuously decreases and roughly halves per kilometre altitude. At 4 km above the tTP, ∼ 3000 observations are available.

The higher moisture values in the ERA5 data become apparent in the vertical profile of the humidity bias (Fig. 5b) that is weakly positive (0.2; 15 %) in the UT and associated with a high standard deviation. This is a result of strong positive and negative bias values, as seen for example in the

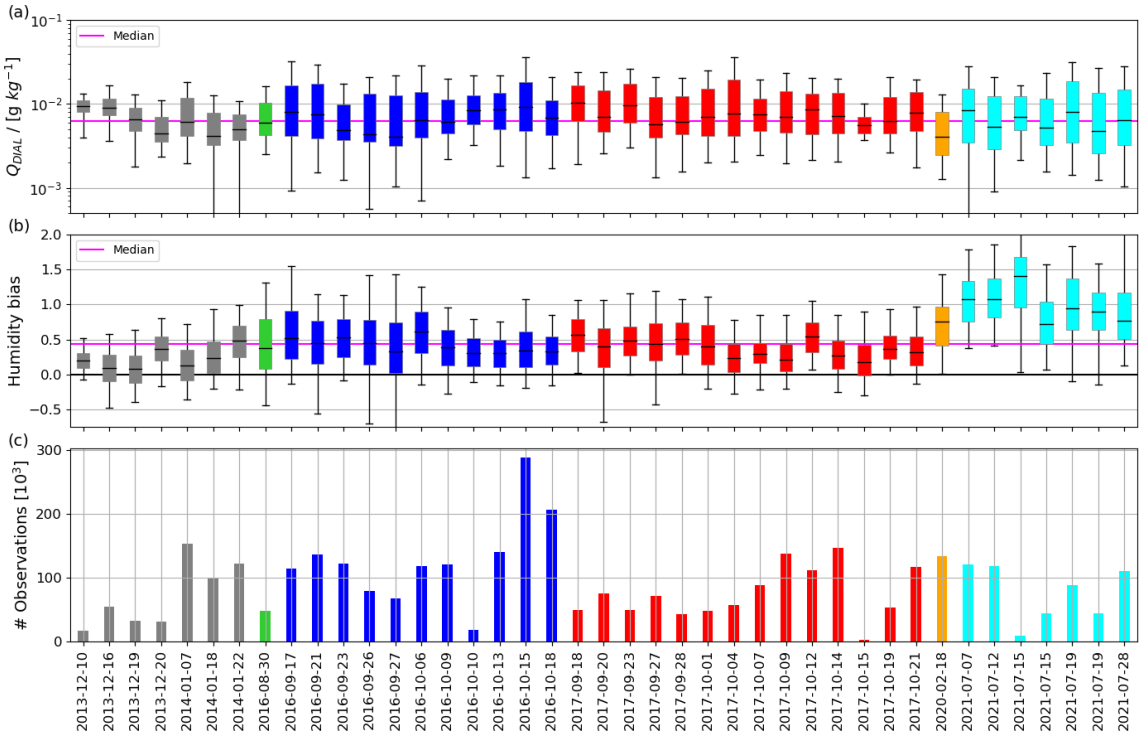

Figure 7. Distributions of **(a)** the observed humidity and **(b)** the humidity bias and **(c)** the number of data points for the layer 0 to 3 km above the tTP. The average observed humidity and bias for all flights is given by the magenta lines in **(a)** and **(b)**. The boxes in **(a)** and **(b)** define the interquartile range located around the median (black), and the whiskers illustrate the 5th and 95th percentile. The different campaigns are colour-coded as in Fig. 1.

case study (Fig. 4c). The weakest bias of 0.2 ($< 15\,\%$) is reached at the tTP level. Above, the vertical moisture gradient is stronger in observations, leading to a significant overestimation of humidity in the LS up to 4 km above tTP. The bias increases to a maximum of 0.63 (55 %) at 1.3 km altitude above the tTP. Beyond, the bias reduces by roughly 0.2 per 500 m up to 4 km altitude above the tTP, where it is approx. 0.2 (15 %). At the highest altitudes (> 4 km above the tTP), a weak and vertically nearly constant bias is observed. At the tropopause as well as above, the standard deviation is significantly reduced compared to the UT. Mean and median profiles of the humidity bias slightly differ, but these differences are very small compared to the magnitude of the bias. The maximum mean and median biases are 0.63 (55 %) and 0.58 (49 %), respectively.

To better illustrate the variability of the water vapour observations in the vertical, Fig. 6 shows the number of data and the mean bias in bins of tropopause-relative altitude and specific humidity. Figure 6a indicates a broader distribution of water vapour observations in the UT compared to the LS. A small number of unusually low humidity values ($< 10^{-2}$ g kg^{-1}) are detected below the tropopause (-4 to -1 km); on the other hand, some data that show high specific humidity ($> 10^{-2}$ g kg^{-1}) are detected at approx. 1–3 km above the tTP. These observations are related to incorrectly assigned tropopause altitudes that were not removed by the applied MPV filtering (see Sect. 2.3.1). However, these remaining outliers are tolerable as they have a negligible impact on the statistics. Throughout the UT, a weak positive bias is detected in bins of highest data availability. At the edges of the distribution, the highest humidity values show a negative bias, while the lowest humidity values stand out due to a positive bias (Fig. 6b). We found that this is related to a narrower distribution of ERA5 humidity compared to the observations (not shown). The low number of observations at the edges should be noted here. In the LS the positive bias is higher and most pronounced up to 3.5 km above the tropopause and at very low specific humidity values. The positive bias reduces towards highest altitudes (> 3 km above the tropopause) of the LS, although the reduced data coverage has to be kept in mind (Fig. 5a).

3.2.2 Synoptic and seasonal variability

In this section, the variability of the LS bias between flights, campaigns, and tropopause altitudes is investigated. Figure 7 shows the observed humidity distribution within a 3 km layer above the tTP, i.e. the area of the strongest LS moist bias. The observed humidity values of all flights range from 1×10^{-2} to 4×10^{-3} g kg^{-1}, and their interquartile range strongly varies between the individual flights, which presumably relates to

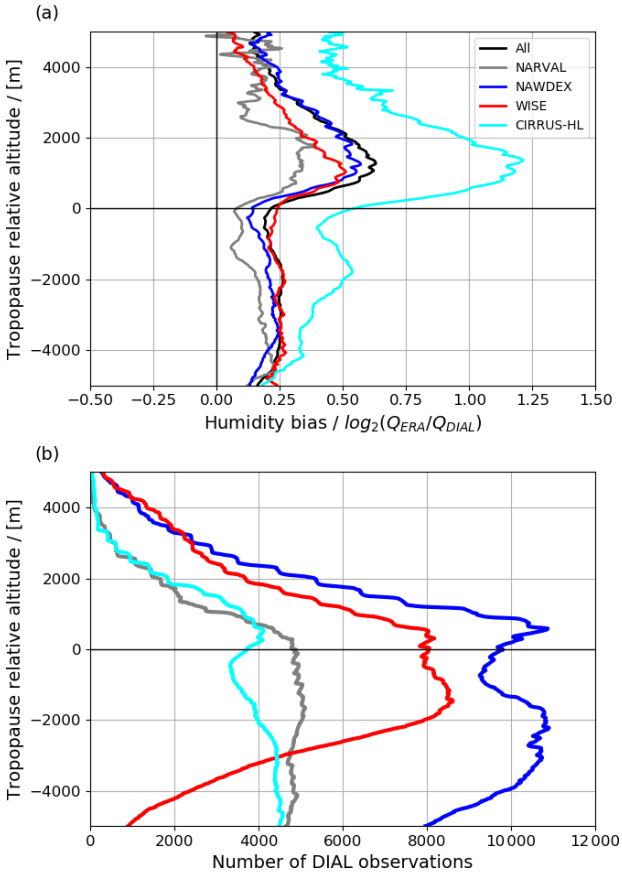

Figure 8. Tropopause-relative vertical profiles of **(a)** humidity bias and **(b)** number of observations for the different campaigns (colour-coded as in Fig. 1). The black line represents the multi-campaign average.

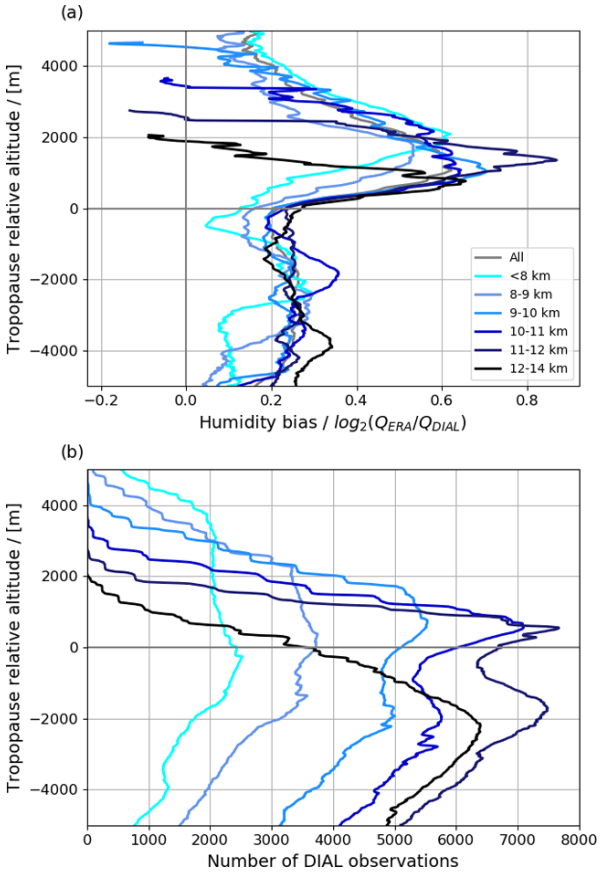

Figure 9. Tropopause-relative profiles of the **(a)** humidity bias and **(b)** number of observations for different intervals of tTP altitudes (colour-coded).

differences in the flight level, the tropopause altitude, and the synoptic situation. During summer (CIRRUS-HL) and autumn campaigns (NAWDEX and WISE), the range of observed humidity is larger compared to the winter campaign (NARVAL). It is furthermore noticeable that intra-campaign variations (i.e. synoptic variability) of observed humidity exceed the seasonal variability. Per flight, the median LS bias (Fig. 7b) varies from 0.2 (15 %) to 1.4 (164 %), but a positive bias is detected for each flight. Whereas the magnitude of the bias shows no obvious correlation with the LS moisture distribution, the moist bias appears to be smaller in winter (NARVAL) compared to autumn (NAWDEX and WISE). Interestingly, the moist bias during the CIRRUS-HL summer campaign is remarkably strong. The number of observations that is available for each flight is strongly variable between a few and several hundred thousand (Fig. 7c).

The average profile of the bias and the number of observations for campaigns with an increased data coverage is shown in Fig. 8. The data availability is very different across the campaigns (Fig. 8b). During NAWDEX and WISE, a large number of observations is present between −5 km below and 4 km above the tTP. CIRRUS-HL provides approximately half as much data at each altitude, except for altitudes beyond 3 km above the tTP, where few data are available. Due to frequent low flight levels during the seven NARVAL flights, only a small number of observations is available beyond 1 km above the tTP. The general structure with a pronounced positive bias, a local minimum at the tropopause, and a decrease towards the highest altitudes is apparent for all campaigns (Fig. 8a), although significant differences can be identified across the campaigns. For the autumn campaigns in 2 successive years (NAWDEX and WISE), a similar shape of the bias is observed across the entire profile. The maximum moist bias is located at approximately the same altitude, and a similar decrease beyond this maximum is observed. However, the magnitude of the LS moist bias is slightly higher for NAWDEX (0.6, 50 %) compared to WISE (0.5, 40 %). During summer (CIRRUS-HL), a stronger moist bias is detected exceeding 1.2 (130 %) at its maximum. Compared to autumn, the summer bias is increased by a factor of 2–3. During winter (NARVAL), the LS moist bias is small (0.3, ∼ 23 %) and not substantially higher than the upper-tropospheric bias, but

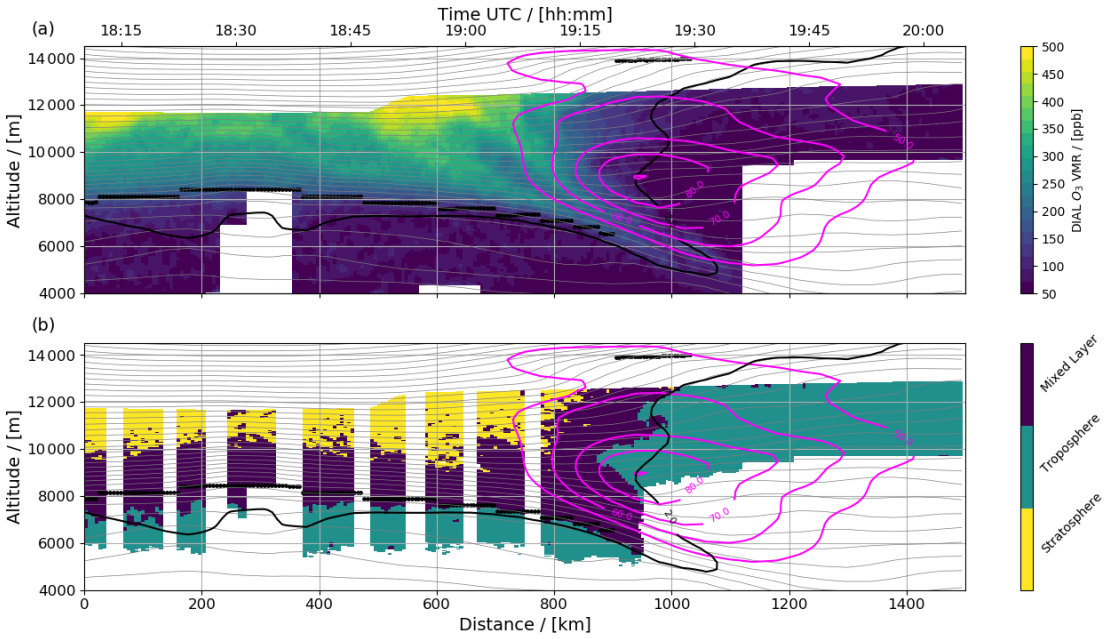

Figure 10. Vertical cross sections as in Fig. 4 but for **(a)** DIAL ozone volume mixing ratio (in ppb) and **(b)** air mass classes derived from water vapour and ozone measurements (for details, see Sect. 3.3).

the limited representativity due to the low number of observations should be noted here.

In addition, we explore whether the observed vertical structure of the moist bias is sensitive to different synoptic situations. For this investigation, the DIAL profiles are classified by their corresponding tTP altitude. Lower tropopauses are typically associated with trough situations and high tropopauses occur above ridges. For each category the corresponding average bias profile and the number of observations is given in Fig. 9. The vertical structure of the bias (Fig. 5b) is reproduced for each tropopause altitude interval. No systematic differences between the bias profiles can be revealed in the UT. Interestingly, each category shows an increased moist bias of comparable magnitude as well as a decrease above, although its vertical position relative to the tTP is different. The maximum bias is located higher for low tropopause altitudes, while profiles with high tTP altitude show a maximum closer to the tTP. For instance, the maximum bias for low tropopauses (< 8 km) is located at 2 km above the tTP, while for the category with highest tropopauses (12–14 km), the maximum value is found at 1 km. The number of data points illustrates that each category exhibits a reasonable number of observations (Fig. 9b).

3.3 The vertical structure of the moist bias related to mixing processes

In the following it is examined to what extent the observed air masses have experienced mixing in their history and whether this is related to the vertical structure of the moist bias. For this purpose, we examine collocated ozone and water vapour observations that were collected during four WISE research flights and that provide a suitable data coverage. First, the observed ozone distribution for the same case study as introduced in Sect. 3.1 is shown in Fig. 10a. Note that the ozone and water vapour observations are given as volume mixing ratios (VMR) in the following. The distribution of ozone is opposite to that of water vapour, with the lowest concentrations ($\mathrm{VMR_{O_3}} < 100$ ppb) in the troposphere and an increase with altitude across the tropopause to $\mathrm{VMR_{O_3}} > 500$ ppb. Note the filamentary structures of increased ozone values in the LS and the ozone-rich air which is transported downward within the tropopause fold (see detailed description in Schäfler et al., 2022).

Following the approach by Schäfler et al. (2021), the collocated water vapour and ozone observations for four WISE flights are illustrated in tracer–tracer (T–T) phase space in Fig. 11, and three classes of observations are identified based on the characteristic distributions (e.g. Pan et al., 2004). First, tropospheric observations are characterized by low $\mathrm{VMR_{O_3}}$ (typically < 100 ppb) and a large spread of $\mathrm{VMR_{H_2O}}$. Second, high $\mathrm{VMR_{O_3}}$ at low $\mathrm{VMR_{H_2O}}$ (< 6.5 ppm or $< 4 \times 10^{-3}$ g kg^{-1}) is assigned to lower-stratospheric air. Third, a class with intermediate chemical characteristics ($\mathrm{VMR_{H_2O}} > 6.5$ ppm and $\mathrm{VMR_{O_3}} > 100$ ppb) is attributed to mixed air masses that experienced mixing between the troposphere and stratosphere. Although Schäfler et al. (2021) suggested a careful selection of the threshold for individual flights, here constant values (see caption of Fig. 11) are used for all four WISE flights, which is sufficient to funda-

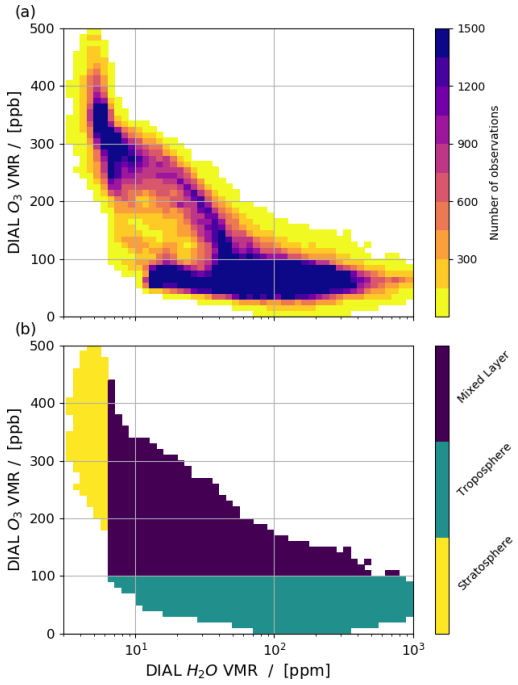

Figure 11. Binned distribution of water vapour and ozone observations in T–T space for four WISE flights coloured by bin-average **(a)** number of DIAL observations and **(b)** type of classified air mass with troposphere ($VMR_{O_3} < 100$ ppb and $VMR_{H_2O} > 6.5$ ppm), mixed air ($VMR_{O_3} > 100$ ppb and $VMR_{H_2O} > 6.5$ ppm), and stratosphere (> 100 ppb VMR_{O_3} and < 6.5 ppm VMR_{H_2O}). Bin sizes are 10 ppb for VMR_{O_3} and 0.05 ppm for $\log_{10}(VMR_{H_2O})$.

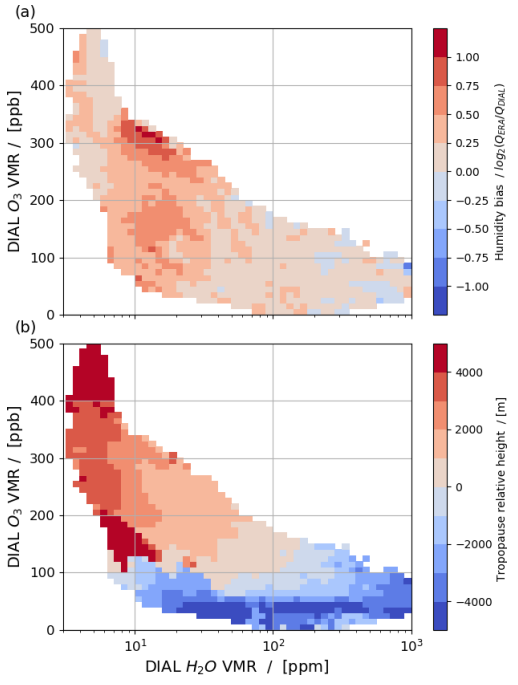

Figure 12. Binned distribution of water vapour and ozone observations in T–T space as in Fig. 11 but coloured by bin-average **(a)** humidity bias and **(b)** tropopause-relative altitude.

mentally identify the three air masses. Sensitivity tests with slightly varied thresholds have shown only a little impact on the distribution of the classes in geometrical space. Such a re-projection of the air mass classification from T–T space to geometrical space with a coherent distribution of the three classes is shown in Fig. 10b. Observations below the tTP are predominantly assigned to tropospheric air, while the uppermost data to the north of the jet stream are classified as stratospheric air. South of the jet stream, where the flight altitude is below the tTP, only tropospheric air is detected. In between the tropospheric and the stratospheric air, the mixed air mass is following the tropopause in a 2–3 km thick layer, which appears to be vertically deeper in the tropopause fold.

For each bin in T–T space, the average humidity bias and the mean tropopause-relative altitude are displayed in Fig. 12. The humidity bias is weak for both tropospheric and stratospheric air (Figs. 12a and 11b), ranging mostly between -0.25 and 0.25. In the mixed air class, the humidity bias is most pronounced (> 0.25), particularly where the VMR_{H_2O} is below 40 ppm. In the tropospheric and stratospheric air, a stronger positive/negative bias is indicated for lower/higher VMR_{H_2O}, which is associated with the sharper humidity distribution in ERA5 (see discussion in Sect. 3.2.1). Figure 12b displays the tropopause-relative height, which is the vertical distance to the tTP, for each bin. Across the mixed air class, an increase of the tropopause-relative altitude is visible, corresponding to a decrease of VMR_{H_2O} and to an increase of VMR_{O_3}. At low VMR_{H_2O} (< 10 ppm) and low VMR_{O_3} (100–200 ppb), the transition of tropopause-relative altitudes is more abrupt, which is related to tTP variability across the jet stream, for example, as visible in the uppermost part of Fig. 10b. When comparing the tropopause-relative height with the distribution of the bias (Fig. 12a), it is noticeable that the average bias is particularly increased between 1 and 3 km, where it ranges from 0.5 (40 %) up to 1.25 (137 %). In contrast, the mean bias is weak beyond 3 km above the tTP and below the tropopause.

The average vertical profile of the moist bias for the WISE flights (Fig. 13a) is similar to the full data set (Fig. 5b) at the tTP and in the LS, i.e. a local minimum is found at the tTP (0.1; 7 %) and a pronounced maximum of 0.62 (54 %) peaking at about 1 km above the tTP. The tropospheric part of the profile, however, is almost constant in the full data set (0.2–0.25) but decreases with increasing altitude in the WISE data (0.4–0.1). Figure 13b shows the relative proportion of the individual air masses at a given tropopause-relative altitude and thus provides information about the connection between the vertical structure of the moist bias and the air mass classes. In the entire UT, the tropospheric air provides the largest contribution of more than 80 % up to 500 m below the tTP. Across the tTP, the proportion of tropospheric air rapidly decreases with altitude in accordance with a rapid growth of the frac-

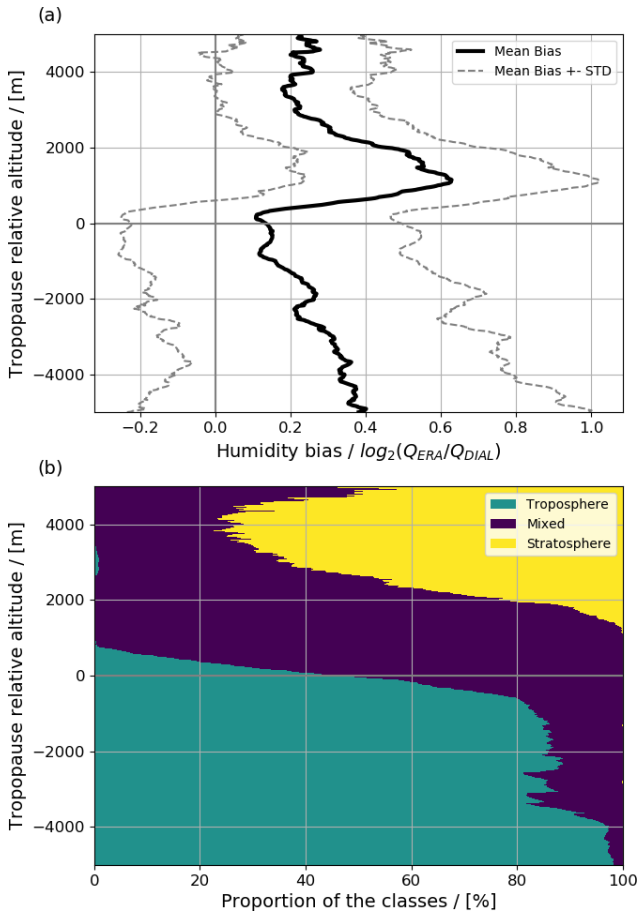

Figure 13. **(a)** Tropopause-relative vertical profile of the mean (thick black line) and standard deviation (dotted thin grey lines) of the humidity bias and **(b)** relative proportion of the individual air mass classes for four WISE flights.

tion of mixed air. This is accompanied by an increase of the moist bias, and the altitude of the largest bias (1–2 km above the tTP) coincides with the maximum relative contribution of the mixed air class ($> 90\,\%$). Above, the relative fraction of stratospheric air grows, while the moist bias reduces and reaches constant values (0.2) at ~ 4 km above the tTP with a 65 %–85 % share of stratospheric air. Please note that contributions of mixed air below the tropopause and at altitudes > 4 km above the tTP may be related to falsely detected tropopause altitudes (see discussion in Sect. 4) or situations of complex tropopause structure (e.g. as shown in the second part of Fig. 10b).

4 Discussion

Recent studies document a lower-stratospheric moist bias in different NWP models (e.g. Kunz et al., 2014; Dyroff et al., 2015; Kaufmann et al., 2018; Woiwode et al., 2020; Bland et al., 2021). We find a comparable moist bias in ERA5 reanalyses based on a comprehensive multi-campaign water vapour lidar data set comprising 41 research flights (six campaigns) and roughly 33 000 vertical profiles obtained in the northern hemispheric mid-latitudes during different seasons. The observations from the surface up to the LS cover 4 orders of magnitude and represent typical mid-latitude data for the individual seasons (e.g. Pan et al., 2000; Randel and Wu, 2010; Zahn et al., 2014; Kunz et al., 2014; Dyroff et al., 2015; Bland et al., 2021). The high data availability around the tropopause makes the data set suitable for an evaluation of NWP fields in the UTLS. Although the number of observations reduces considerably towards the highest altitudes (up to 5 km above the tTP), the data set provides a valuable extension to previous humidity data sets which exhibit increased measurement uncertainties at altitudes larger than 2 km above the tTP (e.g. Bland et al., 2021).

In the troposphere we find strong positive and negative biases of small spatial extent, which are likely related to insufficiently represented tropospheric transport processes, to model errors of tropospheric processes (e.g. clouds), or to the linear interpolation scheme that may have caused increased differences especially in situations of strong horizontal or vertical moisture gradients. The small positive and vertically almost constant mean bias in the UT, which ranges between 0.2 (15 %) and 0.26 (20 %), confirms earlier findings (Dyroff et al., 2015; Bland et al., 2021). It has to be noted that the UT bias is limited to cloud-free scenes, as DIAL humidity profile observations cannot be retrieved inside or below optical thick clouds. In agreement with Bland et al. (2021), a local minimum of the bias (< 0.2, $< 15\,\%$) is found at the tTP. Above the tropopause, our findings confirm a coherent layer of overestimated humidity in ERA5 reanalyses. The magnitude of the maximum bias of 0.63 (55 %) and its altitude of 1.3 km above the tTP are comparable to previous findings for earlier model cycles of the IFS (Dyroff et al., 2015; Kaufmann et al., 2018; Woiwode et al., 2020; Bland et al., 2021), earlier reanalysis versions (Oikonomou and O'Neill, 2006; Kunz et al., 2014), and other evaluated models (Davis et al., 2017; Bland et al., 2021). Above the maximum bias, in a region where recent studies present diverging results (Dyroff et al., 2015; Woiwode et al., 2020; Bland et al., 2021), our analysis reveals a steadily decreasing moist bias that reduces to nearly constant and small positive values comparable to the UT. The independence of measurement error from altitude and humidity concentration allows for a reliable and robust depiction of the bias at the highest altitudes of the UTLS. Furthermore, the magnitude of the LS moist bias exceeds the expected error of the DIAL humidity observations by approx. 1 order of magnitude, which underlines the significance of our results. Please note that Bland et al. (2021) show that tTP altitudes are on average about 200 m higher when derived from ECMWF IFS profiles compared to radiosondes, which may impact tropopause-relative moisture distributions and in turn the bias. As no temperature observations are available, this study relies only on ERA5 tTP altitudes. Assuming a systematic shift by 200 m would reduce

the tropospheric bias, however, the LS moist bias, although slightly weakened, would persist.

In line with findings of Bland et al. (2021), who indicated little sensitivity of the moist bias to various atmospheric conditions but revealed a different depth of the moist bias for trough and ridge situations, low tTP situations (which are typically associated with troughs) exhibit a maximum bias at higher altitudes and a deeper layer of the increased bias compared to high tTP situations. The magnitude of the moist bias is found to be independent of the tropopause altitude. In addition, we detect a pronounced LS moist bias in the summer (> 1.20, $> 130\,\%$) which exceeds the diagnosed autumn bias by a factor of 2–3. So far, such a seasonality has only been suggested in Dyroff et al. (2015). Additional DIAL observations in spring, summer, and winter would be valuable for a more comprehensive study of the seasonality of the vertical bias structure.

For four flights during the WISE campaign, an air mass classification using collocated water vapour and ozone profile data (Schäfler et al., 2021) was applied to separate tropospheric (low ozone and large water vapour mixing ratio), stratospheric (large ozone and low water vapour), and mixed air (intermediate ozone and water vapour). In tropopause-relative coordinates, the vertical structure of the moist bias for the selected cases turned out to be comparable to the multi-campaign LS moist bias, so that these flights are considered to be representative of autumn. We find that the moist bias is increased in the mixed air class representing the ExTL and that the maximum is reached at the altitude where the proportion of mixed air is highest (near 100 %). The decrease of the moist bias above/below is accompanied by a growth of the proportion of stratospheric/tropospheric air. The high correlation in the distribution of the moist bias and the ExTL gives a strong hint at the importance of moisture injection into the LS, either due to numerical diffusion across the tropopause or due to insufficiently modelled transport and mixing processes. As the bias in the ExTL is increased in each of the evaluated WISE flights, we consider systematic uncertainties in the representation of mixing processes to play a key role for the LS moist bias. This is supported by the finding of a deeper bias layer above troughs which are characterized by a thicker ExTL above (e.g. Hoor et al., 2002; Pan et al., 2007). In addition, the maximum bias occurs in summer when cross-tropopause mixing is strongest (Hoor et al., 2002), and, finally, the bias is reduced in stratospheric background humidity at highest altitudes, which are not influenced by mixing processes at the extratropical tropopause. Schäfler et al. (2022) investigate the Lagrangian history of the observed air for the presented WISE case study on 1 October 2017 and find that the ExTL air experienced strong turbulent mixing in the jet stream during 48 h before the observation. They also find that the mixed air (in which we identified the increased bias) shows highly variable origins and transport pathways related to tropospheric weather systems which may be indicative of the relevance of different mixing processes. Additional collocated observations of ozone and water vapour in different seasons, near active mixing process (e.g. convection), or in the Southern Hemisphere where exchange at the polar jet stream is reduced (e.g. Bowman, 1995) could provide valuable information about the relevance of individual mixing processes and their role in forming the moist bias. The presented results suggest that improving the representation of mixing at the tropopause may reduce the humidity bias and be beneficial to improve the modelling of climate and weather. Davis et al. (2017) demonstrate that various reanalyses significantly overestimate LS humidity in the extratropics. The systematic moist bias in ERA5 reanalyses has to be kept in mind for climatological studies using ERA5 humidity fields in the LS.

5 Conclusion

In this study we applied a comprehensive data set of airborne water vapour lidar profiles to investigate the representation of specific humidity in the ERA5 reanalysis across the extratropical UTLS. The main conclusions of this work are summarized below following the three research questions that were raised in the Introduction:

1. *Can the multi-campaign DIAL data set robustly quantify the LS moisture bias in ERA5?*

   The presented DIAL data set with its large number of high-accuracy and high-resolution humidity profiles measured over the North Atlantic and Europe during six research campaigns between 2013–2021 provides a valuable extension to the available observational data sets that were used to determine the lower-stratospheric moist bias. Beside the broad range of observed humidity values (10^{-3} to $10^{1}\,\mathrm{g\,kg^{-1}}$), in particular, the high data availability in the ± 5 km around the tropopause makes the data suitable for the characterization of water vapour in the entire mid-latitude UTLS. The flights that were performed at different times of the year indicate seasonal differences in the observed humidity distributions. As the flights also cover diverse synoptic situations, we consider the data set to be representative of the mid-latitudes. The data set holds the advantage of not being assimilated by NWP and thus allows humidity errors in the ERA5 reanalysis to be evaluated independently.

2. *What is the vertical structure of the LS moist bias in ERA5, particularly at high altitudes?*

   Our analysis demonstrates that a systematic lower-stratospheric moist bias is also present in ECMWF's most recent global reanalysis ERA5. We find that the vertical structure of the bias, which is analysed in tropopause-relative coordinates, is characterized by a weak positive bias in the upper troposphere (15 %–20 %) and a strong overestimation of humidity that reaches a maximum (55 %) at 1.3 km above the thermal

tropopause. Above this maximum, we detect a steady vertical decrease of the moist bias towards a constant small value (15 %) beyond 4 km above the tropopause. The moist bias occurs in coherent and extended regions along the individual lidar cross-sections. The above described unique measurement characteristics of the DIAL data set together with the persistence of the bias structure in different flights and campaigns allow the vertical decline at the highest altitudes to be robustly confirmed. A high similarity for two campaigns conducted in the same region over the North Atlantic in successive years illustrates the persistence of the vertical structure. We find a seasonality of the moist bias with a maximum in summer and a minimum in winter. Lower tropopause altitudes, which are typically related to troughs, exhibit a deeper layer of increased moist bias, while the moist bias over ridges is confined to a shallow layer.

3. *Is the moist bias correlated to the distribution of mixed air masses in the UTLS?*

   For four flights of the DIAL data set, collocated water vapour and ozone profiles are available and used to classify UTLS air masses according to their chemical characteristics into tropospheric, stratospheric, and mixed air. We find the strongest bias at altitudes dominated by the mixed air class, representing the ExTL, while tropospheric or stratospheric air exhibits a smaller bias. From this correlation, we deduce that insufficiently represented mixing processes or numerical diffusion in ERA5 shapes the vertical structure of the lower-stratospheric bias, with the maximum occurring at altitudes that are most frequently affected by exchange processes between the troposphere and the stratosphere. The vertical structure of the moist bias of the entire data set is comparable to the four flights with collocated ozone and water vapour observations. In addition, the deeper bias over troughs which typically feature a deeper ExTL, the maximum moist bias in summer when cross-tropopause mixing is strongest, and the reduced bias at altitudes of constant stratospheric background humidity lead to the conclusion that the findings are applicable to the mid-latitudes in general. In the future, it would be interesting to identify the individual mixing processes that affect the moist bias most and the timescales on which it is formed.

Data availability. The data used in this study are available (upon request from the mission PI) from the HALO database: https://halo-db.pa.op.dlr.de/list/missions (German Aerospace Center, 2021). The ERA5 data were retrieved from the ECMWF Meteorological Archival and Retrieval System (MARS): https://www.ecmwf.int/en/forecasts/access-forecasts/access-archive-datasets (ECMWF, 2022).

Author contributions. KK performed the data analysis, produced the figures, and wrote the manuscript. AS, MWi, MWe, and GCC supported the interpretation of the data, contributed with ideas, and commented on the paper. MWi performed the DIAL data processing.

Competing interests. The contact author has declared that none of the authors has any competing interests.

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

Special issue statement. This article is part of the special issue "WISE: Wave-driven isentropic exchange in the extratropical upper troposphere and lower stratosphere (ACP/AMT/WCD inter-journal SI)". It is not associated with a conference.

Acknowledgements. The authors thank the individual research teams that successfully conducted the field campaigns NARVAL, NARVAL2, NAWDEX, WISE, EUREC^{4}A, and CIRRUS-HL, which enabled us to perform this study. This work was supported by the Transregional Collaborative Research Center SFB/TRR165 "Waves to Weather" (https://www.wavestoweather.de, last access: 5 December 2022), funded by the German Research Foundation (DFG). We further acknowledge the DFG for supporting the HALO missions within the priority programme SPP 1294 "Atmospheric and Earth System Research with HALO" (https://www.halo-spp.de/, last access: 5 December 2022). We are grateful to DLR, who supported this work in the framework of the DLR project "Klimarelevanz von atmosphärischen Spurengasen, Aerosolen und Wolken" (KliSAW). We thank Andreas Dörnbrack for his valuable comments on the manuscript.

Financial support. This research has been supported by the Deutsche Forschungsgemeinschaft (project A3 of the Transregional Collaborative Research Center SFB/TRR 165, "Waves to Weather" TRR 165).

The article processing charges for this open-access publication were covered by the German Aerospace Center (DLR).

Review statement. This paper was edited by Farahnaz Khosrawi and reviewed by two anonymous referees.

References

Bhawar, R., Di Girolamo, P., Summa, D., Flamant, C., Althausen, D., Behrendt, A., Kiemle, C., Bosser, P., Cacciani, M., Champollion, C., Di Iorio, T., Engelmann, R., Herold, C., Müller, D., Pal, S., Wirth, M., and Wulfmeyer, V.: The water

vapour intercomparison effort in the framework of the Convective and Orographically-induced Precipitation Study: airborne-to-ground-based and airborne-to-airborne lidar systems, Q. J. Roy. Meteor. Soc., 137, 325–348, https://doi.org/10.1002/qj.697, 2011.

Birner, T. and Bönisch, H.: Residual circulation trajectories and transit times into the extratropical lowermost stratosphere, Atmos. Chem. Phys., 11, 817–827, https://doi.org/10.5194/acp-11-817-2011, 2011.

Birner, T., Dörnbrack, A., and Schumann, U.: How sharp is the tropopause at midlatitudes?, Geophys. Res. Lett., 29, 45-1–45-4, https://doi.org/10.1029/2002GL015142, 2002.

Bland, J., Gray, S., Methven, J., and Forbes, R.: Characterizing extratropical near-tropopause analysis humidity biases and their radiative effects on temperature forecasts, Q. J. Roy. Meteor. Soc., 140, 3878–3898, https://doi.org/10.1002/qj.4150, 2021.

Brewer, A. W.: Evidence for a world circulation provided by the measurements of helium and water vapor distribution in the stratosphere, Q. J. Roy. Meteor. Soc., 75, 351–363, https://doi.org/10.1002/qj.49707532603, 1949.

Bowman, K. P.: Rossby wave phase speeds and mixing barriers in the stratosphere. Part I: Observations, J. Atmos. Sci., 53, 905–918, https://doi.org/10.1175/1520-0469(1996)053<0905:RWPSAM>2.0.CO;2, 1995.

Chagnon, J. M., Gray, S. L., and Methven, J.: Diabatic processes modifying potential vorticity in a North Atlantic cyclone, Q. J. Roy. Meteor. Soc., 139, 1270–1282, https://doi.org/10.1002/qj.2037, 2013.

Davis, S. M., Hegglin, M. I., Fujiwara, M., Dragani, R., Harada, Y., Kobayashi, C., Long, C., Manney, G. L., Nash, E. R., Potter, G. L., Tegtmeier, S., Wang, T., Wargan, K., and Wright, J. S.: Assessment of upper tropospheric and stratospheric water vapor and ozone in reanalyses as part of S-RIP, Atmos. Chem. Phys., 17, 12743–12778, https://doi.org/10.5194/acp-17-12743-2017, 2017.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrea, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., von de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kallberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553–597, https://https://doi.org/10.1002/qj.828, 2011.

Dessler, A. E. and Sherwood, S. C.: Effect of convection on the summertime extratropical lower stratosphere, J. Geophys. Res., 109, D23301, https://doi.org/10.1029/2004JD005209, 2004.

Diamantakis, M. and Flemming, J.: Global mass fixer algorithms for conservative tracer transport in the ECMWF model, Geosci. Model Dev., 7, 965–979, https://doi.org/10.5194/gmd-7-965-2014, 2014.

Dobson, G. M. B., Brewer, A. W., and Cwilong, B.: The meteorology of the stratosphere, P. Roy. Soc. Lond. A, 185, 144–175, https://doi.org/10.1098/rspa.1946.0010, 1946.

Dyroff, C., Zahn, A., Christner, E., Forbes, R., Tompkins, A. M., and van Velthoven, P. F. J.: Comparison of ECMWF analysis and forecast humidity data with CARIBIC upper troposphere and lower stratosphere observations, Q. J. Roy. Meteor. Soc., 141, 833–844, https://doi.org/10.1002/qj.2400, 2015.

ECMWF: IFS Documentation – Cy41r1: Part III: Dynamics and Numerical Procedures, IFS Documentation, ECMWF, https://www.ecmwf.int/sites/default/files/elibrary/2015/9210-part-iii-dynamics-and-numerical-procedures.pdf (last access: 16 March 2021), 2015.

ECMWF: ECMWF Reanalysis v5 (ERA5), European Centre for Medium-Range Weather Forecasts [data set], https://www.ecmwf.int/en/forecasts/access-forecasts/access-archive-datasets, last access: 29 March 2022.

Ehret, G., Hoinka, K. P., Stein, J., Fix, A., Kiemle, C., and Poberaj, G.: Low stratospheric water vapour measured by an airborne DIAL, J. Geophys. Res.-Atmos., 104, 31351–31359, https://doi.org/10.1029/1999JD900959, 1999.

Esselborn, M., Wirth, M., Fix, A., Tesche, M., and Ehret, G.: Airborne high spectral resolution lidar for measuring aerosol extinction and backscatter coefficients, Appl. Optics, 47, 346–358, https://doi.org/10.1364/AO.47.000346, 2008.

Fix, A., Steinebach, F., Wirth, M., Schäfler, A., and Ehret, G.: Development and application of an airborne differential absorption lidar for the simultaneous measurement of ozone and water vapor profiles in the tropopause region, Appl. Optics, 58, 5892–5900, https://doi.org/10.1364/AO.58.005892, 2019.

Flentje, H., Dörnbrack, A., Fix, A., Ehret, G., and Hólm, E.: Evaluation of ECMWF water vapour fields by airborne differential absorption lidar measurements: a case study between Brazil and Europe, Atmos. Chem. Phys., 7, 5033–5042, https://doi.org/10.5194/acp-7-5033-2007, 2007.

Forster, P. M. F. and Shine, K. P.: Assessing the climate impact of trends in stratospheric water vapor, J. Geophys. Res., 29, 10-1–10-4, https://doi.org/10.1029/2001GL013909, 2002.

Fueglistaler, S., Dessler, A. E., Dunkerton, T. J., Folkins, I., Fu, Q., and Mote, P. W.: Tropical tropopause layer, Rev. Geophys., 47, RG1004, https://doi.org/10.1029/2008RG000267, 2009.

German Aerospace Center: HALO database, German Aerospace Center [data set], https://halo-db.pa.op.dlr.de/list/missions, last access: 4 June 2021.

Gettelman, A., Hoor, P., Pan, L. L., Randel, W. J., Hegglin, M. I., and Birner, T.: The extratropical upper troposphere and lower stratosphere, Rev. Geophys., 49, RG3003, https://doi.org/10.1029/2011RG000355, 2011.

Gray, S., Dunning, C., Methven, J., Masato, G., and Chagnon, J.: Systematic model forecast error in Rossby wave structure, Geophys. Res. Lett., 41, 2979–2987, https://doi.org/10.1002/2014GL059282, 2014.

Groß, S., Esselborn, M., Weinzierl, B., Wirth, M., Fix, A., and Petzold, A.: Aerosol classification by airborne high spectral resolution lidar observations, Atmos. Chem. Phys., 13, 2487–2505, https://doi.org/10.5194/acp-13-2487-2013, 2013.

Haynes, P. and Shuckburgh, E.: Effective diffusivity as a diagnostic of atmospheric transport 2. Troposphere and lower stratosphere, J. Geophys. Res., 105, 22795–22810, https://doi.org/10.1029/2000JD900092, 2000.

Hegglin, M. I., Tegtmeier, S., Anderson, J., Froidevaux, L., Fuller, R., Funke, B., Jones A., Lingenfelser, G., Lumpe, J., Pendlebury, D., Remsberg, E., Rozanov, A., Toohey, M., Urban, J., von Clarmann, T., Walker, K. A., Wang, R., and Weigel, K.: SPARC Data Initiative: Comparison of water vapor climatologies from inter-

national satellite limb sounders, J. Geophys. Res.-Atmos., 118, 11824–11846, https://doi.org/10.1002/jgrd.50752, 2013.

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

Hintsa, E. J., Weinstock, E. M., Dessler, A. E., Anderson, J. G., Loewenstein, M., and Podolske, J. R.: SPADE H_2O measurements and the seasonal cycle of stratospheric water vapor, Geophys. Res. Lett., 21, 2559-2562, https://doi.org/10.1029/94GL01279, 1994.

Holton, J. R., Haynes, P. H., McIntyre, M. E., Douglass, A. R., Rood, R. B., and Pfister L.: Stratosphere-troposphere exchange, Rev. Geophys., 33, 403–439, https://doi.org/10.1029/95RG02097, 1995.

Homeyer, C. R., Pan, L. L., and Barth, M. C.: Transport from convective overshooting of the extratropical tropopause and the role of large-scale lower stratosphere stability, J. Geophys. Res.-Atmos., 119, 2220–2240, https://doi.org/10.1002/2013JD020931, 2014.

Hoor, P., Fischer, H., Lange, L., Lelieveld, J., and Brunner, D.: Seasonal variations of a mixing layer in the lowermost stratosphere as identified by the CO-O_3 correlation from in situ measurements, J. Geophys. Res.-Atmos., 107, 4044, https://doi.org/10.1029/2000JD000289, 2002.

Hoor, P., Wernli, H., Hegglin, M. I., and Bönisch, H.: Transport timescales and tracer properties in the extratropical UTLS, Atmos. Chem. Phys., 10, 7929–7944, https://doi.org/10.5194/acp-10-7929-2010, 2010.

Jiang, J. H., Su, H., Zhai, C. X., Wu, L. T., Minschwaner, K., Molod, A. M., and Tompkins, A. M.: An assessment of upper troposphere and lower stratosphere water vapor in MERRA, MERRA2, and ECMWF reanalyses using Aura MLS observations, J. Geophys. Res.-Atmos., 120, 11468–11485, https://doi.org/10.1002/2015JD023752, 2015.

Kaufmann, S., Voigt, C., Heller, R., Jurkat-Witschas, T., Krämer, M., Rolf, C., Zöger, M., Giez, A., Buchholz, B., Ebert, V., Thornberry, T., and Schumann, U.: Intercomparison of midlatitude tropospheric and lower-stratospheric water vapor measurements and comparison to ECMWF humidity data, Atmos. Chem. Phys., 18, 16729–16745, https://doi.org/10.5194/acp-18-16729-2018, 2018.

Khosrawi, F., Lossow, S., Stiller, G. P., Rosenlof, K. H., Urban, J., Burrows, J. P., Damadeo, R. P., Eriksson, P., García-Comas, M., Gille, J. C., Kasai, Y., Kiefer, M., Nedoluha, G. E., Noël, S., Raspollini, P., Read, W. G., Rozanov, A., Sioris, C. E., Walker, K. A., and Weigel, K.: The SPARC water vapour assessment II: comparison of stratospheric and lower mesospheric water vapour time series observed from satellites, Atmos. Meas. Tech., 11, 4435–4463, https://doi.org/10.5194/amt-11-4435-2018, 2018.

Kiemle, C., Wirth, M., Fix, A., Ehret, G., Schumann, U., Gardiner, T., Schiller, C., Sitnikov, N., and Stiller, G.: First airborne water vapor lidar measurements in the tropical upper troposphere and mid-latitudes lower stratosphere: accuracy evaluation and intercomparisons with other instruments, Atmos. Chem. Phys., 8, 5245–5261, https://doi.org/10.5194/acp-8-5245-2008, 2008.

Kiemle, C., Schäfler, A., and Voigt, C.: Detection and Analysis of Water Vapor Transport, in: Atmospheric Physics: Background – Methods – Trends, edited by: Schumann, U., Springer-Verlag, Berlin, 169–184, https://doi.org/10.1007/978-3-642-30183-4_11, 2012.

Klepp, C., Ament, F., Bakan, S., Hirsch, L., and Stevens, B.: The Next-generation Aircraft Remote sensing for VALidation studies (NARVAL) Campaign flight reports using the research aircraft HALO (The NARVAL Campaign Report), Berichte zur Erdsystemforschung/Max-Planck-Institut für Meteorologie, 164, https://pure.mpg.de/rest/items/item_2129055/component/file_2129139/content (last access: 30 June 2022), 2014.

Krautstrunk, M. and Giez, A.: The transition from FALCON to HALO era airborne atmospheric research, in: Atmospheric Physics: Background – Methods – Trends, edited by: Schumann, U., Springer-Verlag, Berlin, 609–624, https://doi.org/10.1007/978-3-642-30183-4_37, 2012.

Kunkel, D., Hoor, P., Kaluza, T., Ungermann, J., Kluschat, B., Giez, A., Lachnitt, H.-C., Kaufmann, M., and Riese, M.: Evidence of small-scale quasi-isentropic mixing in ridges of extratropical baroclinic waves, Atmos. Chem. Phys., 19, 12607–12630, https://doi.org/10.5194/acp-19-12607-2019, 2019.

Kunz, A., Spelten, N., Konopka, P., Müller, R., Forbes, R. M., and Wernli, H.: Comparison of Fast In situ Stratospheric Hygrometer (FISH) measurements of water vapor in the upper troposphere and lower stratosphere (UTLS) with ECMWF (re)analysis data, Atmos. Chem. Phys., 14, 10803–10822, https://doi.org/10.5194/acp-14-10803-2014, 2014.

Martius, O., Schwierz, C., and Davies, H. C.: Tropopause-Level Waveguides, J. Atmos. Sci., 67, 866–879, https://doi.org/10.1175/2009JAS2995.1, 2010.

Oikonomou, E. K. and O'Neill, A.: Evaluation of ozone and water vapor fields from the ECMWF reanalysis ERA-40 during 1991–1999 in comparison with UARS satellite and MOZAIC aircraft observations, J. Geophys. Res., 111, D14109, https://doi.org/10.1029/2004JD005341, 2006.

Pan, L., Hintsa, E., Stone, E., Weinstock, E., and Randel, W.: The seasonal cycle of water vapor and saturation vapor mixing ratio in the extratropical lowermost stratosphere, J. Geophys. Res., 105, 26519–26530, https://doi.org/10.1029/2000JD900401, 2000.

Pan, L. L., Randel, W. J., Gary, B. L., Mahoney, M. J., and Hintsa, E. J.: Definitions and sharpness of the extratropical tropopause: A trace gas perspective, J. Geophys. Res.-Atmos., 109, D23103, https://doi.org/10.1029/2004JD004982, 2004.

Pan, L. L., Bowman, K. P., Shapiro, M., Randel, W. J., Gao, R. S., Campos, T., Davis, C., Schauffler, S., Ridley, B. A., Wei, J. C., and Barnet, C.: Chemical behavior of the tropopause observed during the Stratosphere-Troposphere Analyses of Regional Transport experiment, J. Geophys. Res.-Atmos., 112, D18110, https://doi.org/10.1029/2007JD008645, 2007.

Randel, W. J. and Wu, F.: The polar summer tropopause inversion layer, J. Atmos. Sci., 67, 2572–2581, https://doi.org/10.1175/2010JAS3430.1, 2010.

Randel, W. J., Wu, F., and Forster, P.: The extratropical tropopause inversion layer: Global observations with GPS data, and a radiative forcing mechanism, J. Atmos. Sci., 64, 4489–4496, https://doi.org/10.1175/2007JAS2412.1, 2007.

Riese, M., Ploeger, F., Rap, A., Vogel, B., Konopka, P., Dameris, M., and Forster, P.: Impact of uncertainties in atmospheric mixing on simulated UTLS composition and related radiative effects, J. Geophys. Res.-Atmos., 117, D16305, https://doi.org/10.1029/2012JD017751, 2012.

Schäfler, A., Dörnbrack, A., Kiemle, C., Rahm, S., and Wirth, M.: Tropospheric water vapour transport as determined from airborne lidar measurements, J. Atmos. Ocean. Tech., 27, 2017–2030, https://doi.org/10.1175/2010JTECHA1418.1, 2010.

Schäfler, A., Craig, G., Wernli, H., Arbogast, P., Doyle, J. D., McTaggart-Cowan, R., Methven, J., Rivière, G., Ament, F., Boettcher, M., Bramberger, M., Cazenave, Q., Cotton, R., Crewell, S., Delanoë, J., Dörnbrack, A., Ehrlich, A., Ewald, F., Fix, A., Grams, C. M., Gray, S. L., Grob, H., Groß, S., Hagen, M., Harvey, B., Hirsch, L., Jacob, M., Kölling, T., Konow, H., Lemmerz, C., Lux, O., Magnusson, L., Mayer, B., Mech, M., Moore, R., Pelon, J., Quinting, J., Rahm, S., Rapp, M., Rautenhaus, M., Reitebuch, O., Reynolds, C. A., Sodemann, H., Spengler, T., Vaughan, G., Wendisch, M., Wirth, M., Witschas, B., Wolf, K., and Zinner, T.: The North Atlantic Waveguide and Downstream Impact Experiment, B. Am. Meteorol. Soc., 99, 1607–1637, https://doi.org/10.1175/BAMS-D-17-0003.1, 2018.

Schäfler, A., Harvey, B., Methven, J., Doyle, J. D., Rahm, S., Reitebuch, O., Weiler, F., and Witschas, B.: Observation of jet stream winds during NAWDEX and characterization of systematic meteorological analysis error, Mon. Weather Rev., 148, 2889–2907, https://doi.org/10.1175/MWR-D-19-0229.1, 2020.

Schäfler, A., Fix, A., and Wirth, M.: Mixing at the extratropical tropopause as characterized by collocated airborne H_2O and O_3 lidar observations, Atmos. Chem. Phys., 21, 5217–5234, https://doi.org/10.5194/acp-21-5217-2021, 2021.

Schäfler, A., Sprenger, M., Wernli, H., Fix, A., and Wirth, M.: Case study on the influence of synoptic-scale processes on the paired H_2O-O_3 distribution in the UTLS across a North Atlantic jet stream, Atmos. Chem. Phys. Discuss. [preprint], https://doi.org/10.5194/acp-2022-692, in review, 2022.

Shapiro, M. A.: Turbulent Mixing within Tropopause Folds as a Mechanism for the Exchange of Chemical Constituents between the Stratosphere and Troposphere, J. Atmos. Sci., 37, 994–1004, https://doi.org/10.1175/1520-0469(1980)037<0994:TMWTFA>2.0.CO;2, 1980.

Shapiro, M. A., Wernli, H., Bao, J., Methven, J., Zou, X., Doyle, J., Holt, T., Donall-Grell, E., and Neiman P.: A Planetary-Scale to Mesoscale Perspective of the Life Cycles of Extratropical Cyclones: The Bridge between Theory and Observations, in: The Life Cycles of Extratropical Cyclones, edited by: Shapiro, M. A. and Grønås, S., American Meteorological Society, Boston, MA, https://doi.org/10.1007/978-1-935704-09-6_14, 1999.

Shepherd, T. G., Polichtchouk, I., Hogan, R. J., and Simmons, A. J.: Report on Stratosphere Task Force, ECMWF Technical Memorandum 824, https://doi.org/10.21957/0vkp0t1xx, 2018.

Stenke, A., Grewe, V., and Ponater, M.: Lagrangian transport of water vapor and cloud water in the ECHAM4 GCM and its impact on the cold bias, Clim. Dynam., 31, 491–506, https://doi.org/10.1007/s00382-007-0347-5, 2008.

Stevens, B., Ament, F., Bony, S., Crewell, S., Ewald, F., Gross, S., Hansen, A., Hirsch, L., Jacob, M., Kölling, T., Konow, H., Mayer, B., Wendisch, M., Wirth, M., Wolf, K., Bakan, S., Bauer-Pfundstein, M., Brueck, M., Delanoë, J., Ehrlich, A., Farrell, D., Forde, M., Gödde, F., Grob, H., Hagen, M., Jäkel, E., Jansen, F., Klepp, C., Klingebiel, M., Mech, M., Peters, G., Rapp, M., Wing, A. A., and Zinner, T.: A High-Altitude Long-Range Aircraft Configured as a Cloud Observatory: The NARVAL Expeditions, B. Am. Meteorol. Soc., 100, 1061–1077, https://doi.org/10.1175/BAMS-D-18-0198.1, 2019.

Stevens, B., Bony, S., Farrell, D., Ament, F., Blyth, A., Fairall, C., Karstensen, J., Quinn, P. K., Speich, S., Acquistapace, C., Aemisegger, F., Albright, A. L., Bellenger, H., Bodenschatz, E., Caesar, K.-A., Chewitt-Lucas, R., de Boer, G., Delanoë, J., Denby, L., Ewald, F., Fildier, B., Forde, M., George, G., Gross, S., Hagen, M., Hausold, A., Heywood, K. J., Hirsch, L., Jacob, M., Jansen, F., Kinne, S., Klocke, D., Kölling, T., Konow, H., Lothon, M., Mohr, W., Naumann, A. K., Nuijens, L., Olivier, L., Pincus, R., Pöhlker, M., Reverdin, G., Roberts, G., Schnitt, S., Schulz, H., Siebesma, A. P., Stephan, C. C., Sullivan, P., Touzé-Peiffer, L., Vial, J., Vogel, R., Zuidema, P., Alexander, N., Alves, L., Arixi, S., Asmath, H., Bagheri, G., Baier, K., Bailey, A., Baranowski, D., Baron, A., Barrau, S., Barrett, P. A., Batier, F., Behrendt, A., Bendinger, A., Beucher, F., Bigorre, S., Blades, E., Blossey, P., Bock, O., Böing, S., Bosser, P., Bourras, D., Bouruet-Aubertot, P., Bower, K., Branellec, P., Branger, H., Brennek, M., Brewer, A., Brilouet , P.-E., Brügmann, B., Buehler, S. A., Burke, E., Burton, R., Calmer, R., Canonici, J.-C., Carton, X., Cato Jr., G., Charles, J. A., Chazette, P., Chen, Y., Chilinski, M. T., Choularton, T., Chuang, P., Clarke, S., Coe, H., Cornet, C., Coutris, P., Couvreux, F., Crewell, S., Cronin, T., Cui, Z., Cuypers, Y., Daley, A., Damerell, G. M., Dauhut, T., Deneke, H., Desbios, J.-P., Dörner, S., Donner, S., Douet, V., Drushka, K., Dütsch, M., Ehrlich, A., Emanuel, K., Emmanouilidis, A., Etienne, J.-C., Etienne-Leblanc, S., Faure, G., Feingold, G., Ferrero, L., Fix, A., Flamant, C., Flatau, P. J., Foltz, G. R., Forster, L., Furtuna, I., Gadian, A., Galewsky, J., Gallagher, M., Gallimore, P., Gaston, C., Gentemann, C., Geyskens, N., Giez, A., Gollop, J., Gouirand, I., Gourbeyre, C., de Graaf, D., de Groot, G. E., Grosz, R., Güttler, J., Gutleben, M., Hall, K., Harris, G., Helfer, K. C., Henze, D., Herbert, C., Holanda, B., Ibanez-Landeta, A., Intrieri, J., Iyer, S., Julien, F., Kalesse, H., Kazil, J., Kellman, A., Kidane, A. T., Kirchner, U., Klingebiel, M., Körner, M., Kremper, L. A., Kretzschmar, J., Krüger, O., Kumala, W., Kurz, A., L'Hégaret, P., Labaste, M., Lachlan-Cope, T., Laing, A., Landschützer, P., Lang, T., Lange, D., Lange, I., Laplace, C., Lavik, G., Laxenaire, R., Le Bihan, C., Leandro, M., Lefevre, N., Lena, M., Lenschow, D., Li, Q., Lloyd, G., Los, S., Losi, N., Lovell, O., Luneau, C., Makuch, P., Malinowski, S., Manta, G., Marinou, E., Marsden, N., Masson, S., Maury, N., Mayer, B., Mayers-Als, M., Mazel, C., McGeary, W., McWilliams, J. C., Mech, M., Mehlmann, M., Meroni, A. N., Mieslinger, T., Minikin, A., Minnett, P., Möller, G., Morfa Avalos, Y., Muller, C., Musat, I., Napoli, A., Neuberger, A., Noisel, C., Noone, D., Nordsiek, F., Nowak, J. L., Oswald, L., Parker, D. J., Peck, C., Person, R., Philippi, M., Plueddemann, A., Pöhlker, C., Pörtge, V., Pöschl, U., Pologne, L., Posyniak, M., Prange, M., Quiñones Meléndez, E., Radtke, J., Ramage, K., Reimann, J., Renault, L., Reus, K., Reyes, A., Ribbe, J., Ringel, M., Ritschel, M., Rocha, C. B.,

Rochetin, N., Röttenbacher, J., Rollo, C., Royer, H., Sadoulet, P., Saffin, L., Sandiford, S., Sandu, I., Schäfer, M., Schemann, V., Schirmacher, I., Schlenczek, O., Schmidt, J., Schröder, M., Schwarzenboeck, A., Sealy, A., Senff, C. J., Serikov, I., Shohan, S., Siddle, E., Smirnov, A., Späth, F., Spooner, B., Stolla, M. K., Szkółka, W., de Szoeke, S. P., Tarot, S., Tetoni, E., Thompson, E., Thomson, J., Tomassini, L., Totems, J., Ubele, A. A., Villiger, L., von Arx, J., Wagner, T., Walther, A., Webber, B., Wendisch, M., Whitehall, S., Wiltshire, A., Wing, A. A., Wirth, M., Wiskandt, J., Wolf, K., Worbes, L., Wright, E., Wulfmeyer, V., Young, S., Zhang, C., Zhang, D., Ziemen, F., Zinner, T., and Zöger, M.: EUREC4A, Earth Syst. Sci. Data, 13, 4067–4119, https://doi.org/10.5194/essd-13-4067-2021, 2021.

Stohl, A., Bonasoni, P., Cristofanelli, P., Collins, W., Feichter, J., Frank, A., Forster, C., Gerasopoulos, E., Gaggeler, H., James, P., Kentarchos, T., Kromp-Kolb, H., Kruger, B., Land, C., Meloen, J., Papayannis, A., Priller, A., Seibert, P., Sprenger, M., Roelofs, G. J., Scheell, H., E. Schnabel, C., Siegmund, P., Tobler, L., Trickl, T., Wernli, H., Wirth, V., Zanis, P., and Zerefos, C.: Stratosphere-troposphere exchange: A review, and what we have learned from STACCATO, J. Geophys. Res.-Atmos., 108, 8516, https://doi.org/10.1029/2002JD002490, 2003.

Trickl, T., Vogelmann, H., Fix, A., Schäfler, A., Wirth, M., Calpini, B., Levrat, G., Romanens, G., Apituley, A., Wilson, K. M., Begbie, R., Reichardt, J., Vömel, H., and Sprenger, M.: How stratospheric are deep stratospheric intrusions? LUAMI 2008, Atmos. Chem. Phys., 16, 8791–8815, https://doi.org/10.5194/acp-16-8791-2016, 2016.

Wirth, M., Fix, A., Mahnke, P., Schwarzer, H., Schrandt, F., and Ehret, G.: The airborne multi-wavelength water vapour differential absorption lidar WALES: system design and performance, Appl. Phys. B, 96, 201–213, https://doi.org/10.1007/s00340-009-3365-7, 2009.

WMO (World Meteorological Organization): Meteorology – A three-dimensional science, WMO Bull., 6, 134–138, 1957.

Woiwode, W., Dörnbrack, A., Polichtchouk, I., Johansson, S., Harvey, B., Höpfner, M., Ungermann, J., and Friedl-Vallon, F.: Technical note: Lowermost-stratosphere moist bias in ECMWF IFS model diagnosed from airborne GLORIA observations during winter–spring 2016, Atmos. Chem. Phys., 20, 15379–15387, https://doi.org/10.5194/acp-20-15379-2020, 2020.

Zahn, A., Christner, E., van Velthoven, P. F. J., Rauthe-Schoch, A., and Brenninkmeijer, C. A. M.: Processes controlling water vapor in the upper troposphere/lowermost stratosphere: An analysis of 8 years of monthly measurements by the IAGOS-CARIBIC observatory, J. Geophys. Res.-Atmos., 119, 11505–11525, https://doi.org/10.1002/2014JD021687, 2014.