Influence of radiosonde observations on the sharpness and altitude of the midlatitude tropopause in the ECMWF IFS

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

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

Correspondence to: Konstantin Krueger (konstantin.krueger@dlr.de)

Abstract

Initial conditions of current numerical weather prediction systems insufficiently represent the sharp vertical gradients across the midlatitude tropopause. Data assimilation may provide a means to improve the tropopause structure by correcting the erroneous background forecast towards the observations. In this paper, the influence of assimilating radiosonde observations on the tropopause structure, i.e., the sharpness and altitude, is investigated in the ECMWF IFS. We evaluate 9729 midlatitude radiosondes launched during one month in autumn 2016. About 500 of these radiosondes, launched on request during the North Atlantic Waveguide Downstream impact Experiment (NAWDEX) field campaign, were used to set up an observing system experiment (OSE) that ~~. The OSE~~ comprises two ~~cycled~~ assimilation forecast experiments, one run with and one without the non-operational soundings. The influence on the tropopause is assessed in a statistical, tropopause-relative evaluation of observation departures of temperature, static stability (N^2), wind speed and wind shear from the background forecast and the analysis. T~~he background t~~emperature is overestimated by the background at the tropopause (warm bias, ~1 K) and underestimated in the lower stratosphere (cold bias, –0.3 K) leading to an underestimation of the abrupt ~~vertical~~ increase of N^2 at the tropopause. ~~We show that t~~The increments (differences of analysis and background) reduce these background biases and improve the tropopause sharpness. Profiles with sharper tropopause exhibit stronger background biases ~~and, in turn,~~but also an increased positive influence of the observations on temperature and N^2 in the analysis. Wind speed is underestimated in the background, especially in the upper troposphere (~1 m s^{-1}), but the assimilation improves the wind profile. For the strongest winds the background bias is roughly halved. The positive influence on the analysis wind ~~distribution~~ profile is associated with an improved vertical distribution of wind shear, particularly in the lower stratosphere. ~~is associated with an increase of vertical wind speed shear, which is underestimated above the tropopause. In addition to the tropopause sharpening,~~ Wwe furthermore ~~-~~detect a shift of the analysis tropopause altitude towards the observations. The ~~comparison~~ evaluation of the OSE ~~runs~~highlights that ~~the main contribution to~~ the diagnosed tropopause sharpening can be primarily attributed to the radiosondes. This study shows that data assimilation improves wind and temperature gradients across the

30 tropopause, but the sharpening is small compared to the model biases. Hence, the analysis still systematically underestimates the tropopause sharpness which may negatively impact weather and climate forecasts.

1 Introduction

The extratropical tropopause is the physical boundary that separates the well-mixed upper troposphere (UT) from the stably stratified lower stratosphere (LS) (e.g., Gettelman et al., 2011). The transition from the UT to the LS is characterized by sharp

35 vertical gradients of temperature, humidity and wind and the strength of these gradients determines the sharpness and altitude of the tropopause. In the UT, the average temperature decreases with altitude towards a minimum at the tropopause. Above the tropopause, a ~2 km thick temperature inversion is typically followed by a nearly isothermal temperature in the LS. This temperature distribution leads to a rapid increase of the squared static stability (N^2) from low values (1×10^{-4} s^{-2}) in the UT to high values (4×10^{-4} s^{-2}) in the lowermost 2–3 km of the LS referred to as the tropopause inversion layer (TIL; Birner et al.,

40 2002). The N^2 maximum above the tropopause (within the TIL) is used as a metric for the tropopause sharpness (e.g., Haualand and Spengler 2021; Boljka and Birner, 2022). The TIL acts as a barrier for vertical transport leading to sharp gradients of trace species, across the tropopause, e.g., of specific humidity (Krüger et al., 2022). The vertical distribution of wind in the midlatitude UTLS is highly variable, but on average wind speed linearly increases with altitude in the troposphere towards a maximum just below the tropopause (e.g., Birner et al., 2002; Birner, 2006; Schäfler et al., 2020).

45 Above, wind speed rapidly decreases in the LS associated with an increased magnitude of vertical shear of the horizontal wind (Birner, 2006; Schäfler et al., 2020).

Temperature and wind gradients directly determine the potential vorticity (PV) distribution. The strong meridional PV gradient near the tropopause acts as a waveguide for Rossby waves (Schwierz et al., 2004; Martius et al., 2010) and, in turn,

50 impacts downstream weather development in the midlatitudes (Harvey et al., 2018). Thus, an accurate representation of the sharp cross-tropopause gradients in the initial conditions may be of high importance for numerical weather prediction (NWP) models. However, forecast PV gradients rapidly decline within short (12–24 h) lead times (Gray et al., 2014; Lavers et al., 2023) which is attributed to a smoothing effect of the advection scheme that dominates sharpening effects of parameterized processes; such as radiative cooling driven by water vapor, microphysics, and turbulent mixing (Saffin et al.,

55 2017). The weakening PV gradients are likely associated with background forecast errors of temperature, humidity and wind at the tropopause, which may affect the quality of the analysis.

At the tropopause, Bland et al. (2021) found a warm bias (few tenths of K) in analyses of the European Centre for Medium-Range Weather Forecast's (ECMWF) Integrated Forecasting System (IFS). The presence of a warm bias in IFS short-range forecasts and analyses near the tropopause was also indicated in earlier studies (Bonavita, 2014; Ingleby et al., 2016). In the

60 LS, a moist bias (e.g., Krüger et al., 2022) leads to a cold bias at altitudes between 0.5 and 2 km above the tropopause in the IFS (Bland et al., 2021). Schäfler et al. (2020) indicated a systematic underestimation of jet stream wind maxima and showed

2

large wind errors of up to 10 m s^{-1} for individual cases in IFS short~~-~~–range forecasts and analyses. Lavers et al. (2023) detected a vertically increasing slow wind bias ~~in IFS background~~ in the troposphere in the IFS background (up to roughly 0.6 m s^{-1}). ~~This wind speed underestimation in models is in line with Birner et al. (2002) who found notably underestimated UT wind maxima and vertical wind shear in the ERA-15 reanalysis.~~In agreement with the wind errors, ~~q~~Quantitative assessments of the magnitude of vertical wind shear at the tropopause revealed an underestimation by a factor of 2–5 (Houchi et al., 2010; Schäfler et al., 2020).

Data assimilation (DA) has shown a positive influence on the analysis in the UTLS, i.e., a reduction of the short~~-~~–range forecast errors of temperature (e.g., Radn*ó*oti et al., 2010; Bonavita, 2014) and wind (e.g., Weissman and Cardinali, 2007; Weissmann et al., 2012; Lavers et al., 2023; Martin et al., 2023). Two dedicated studies elaborated the influence of DA on the tropopause sharpness. Birner et al. (2006) investigated the role of satellite DA in the Canadian Middle Atmosphere Model (CMAM) providing a vertical resolution of 1 km near the tropopause and used a 3DVAR assimilation scheme. A decrease of N$^2_{max}$ in an experiment with assimilated satellite observations compared to a model run without suggested that satellite DA smears out the gradients near the tropopause. A more recent study by Pilch Kedziersky et al. (2016) analysed the influence on the tropopause sharpness at the positions of GPS radio occultation (GPS–RO) observations in ECMWF's ERA–Interim reanalysis and IFS analysis, using 4DVAR (e.g., Rabier et al., 2000) and a vertical resolution of ~500 m at the tropopause. The detected increase of N^2 in an ~1 km thick layer just above the tropopause and a decrease of N^2 above and below this layer is corresponding to a tropopause sharpening, which was attributed to the assimilation of GPS–RO data. Both studies, which show different effects of DA on the tropopause sharpness, differ in terms of the applied methods to diagnose the influence, the used observation type, the spatial resolution and the DA schemes. ~~Hence, no definitive conclusion can be drawn as to whether DA sharpens or smoothens the tropopause. Additionally, it~~ It should also be noted that both studies are based on variational DA schemes without a flow-dependent estimate of the error covariance matrix (***B***). Flow-dependent estimates of B as they are nowadays used in the hybrid DA scheme of ECMWF are expected to lead to more accurate increment structures and therefore a better representation of sharp gradients.

~~In this study, we investigate radiosonde~~ Radiosondes ~~profiles, which~~ provide highly resolved and accurate profiles of temperature and wind components (e.g., Vaisala, 2017), and thus are suitable to resolve the sharp vertical gradients at the tropopause. The measured quantities are directly assimilated and, although they only account for a small proportion (about 2 %) of the total assimilated meteorological information, they contribute to a 5 % reduction in 24-h forecast error in the ECMWF IFS in a statistical sense (Pauley and Ingleby, 2022). In addition, radiosondes serve as anchor observations for the variational bias correction e.g., for satellite observations, highlighting their important role for DA (Cucurull and Anthes, 2014). The impact of individual observation capabilities such as radiosondes is typically assessed by performing observing system experiments (OSEs; e.g., Bonavita, 2014), e.g., during special observation periods related to field campaigns (e.g., Weissmann et al., 2012; Schindler et al., 2020; Borne et al., 2023). The DA impact can be studied either in model space by using 3D gridded model output or in observation space, which is the 4DVAR model output (observations and departures) representative for the position and time of the assimilated observation. The latter method has the advantage that a comparison of the

3

observations and departures allows the influence of individual measurement types and parameters in the NWP system to be evaluated.

In this study, we further address the question whether DA sharpens or smoothens the near-tropopause gradients. The aim scope of this study is to quantify the change of the tropopause structure from the first-guess to the analysis and to relate it to the assimilation of radiosondes. For this purpose, we make use of the 1-month campaign period of the North Atlantic Waveguide Downstream impact EXperiment (NAWDEX; Schäfler et al., 2018) in autumn 2016 during which 9729 radiosonde profiles in a region between eastern North America and Europe were assimilated. 497 of the 9729 radiosondes were non-operationally launched and applied intheir impact was studied an OSE, which consists of two cycled IFS runs, one with and one without the additional observations (Schindler et al., 2020). The statistical evaluation is performed in a tropopause-relative framework which is mandatory to preserve the outlined sharp gradients in the UTLS when averaging profiles with different tropopause altitudes (Birner, 2006). As no humidity data at and above the tropopause is assimilated (Bland et al., 2021), we restrict the analysis to profile observations of temperature and wind. We address the following specific research questions:

1. How is tropopause sharpness represented in background forecasts and what is the influence of DA on the analysis? Does the diagnosed temperature and wind influence depend on the tropopause structure and vary in different dynamic situations?
2. Does the influence on the temperature profile affect the tropopause altitude?
3. Can the diagnosed influence be attributed to the assimilated radiosondes or do other observations also affect the tropopause structure?

2 Data and Methods

2.1 Description of the data set and the OSE

In this study, we analyse about 9200 radiosonde profiles (Fig. 1a) that were routinely measured at 581 sites covering a wide area between North America and Europe from the subtropics to high latitudes (30°N–85°N, 95°W–30°E) during a 1one-month autumn period (17 September–18 October 2016). The majority of these observations (96 %) were performed at 200 land-based stations while a minor share (4 %) are ship-based observations at 381 variable positions across the North Atlantic. In addition to the routine profile observations, about 500 extra radiosondes were launched in the course of the NAWDEX field campaign (Fig. 1b), which had the aim to better explore the influence of diabatic processes on the polar jet and weather downstream (Schäfler et al., 2018). Over Europe the extra, on-demand radiosondes were released in a variety of synoptic situations, for instance, in diabatically active warm conveyor belt flows associated with cyclones or in upper-level ridges associated with blocking situations. Six stations over Canada, upstream of the NAWDEX operation region, released two additional radiosondes per day. In addition to the radiosonde observations, more than 700 dropsondes were released from research aircrafts during the NAWDEX period (mostly in the subtropical and tropical west Atlantic, see Schindler et al., 2020). Due to the low data coverage of the dropsondes above and at the tropopause related to the limited flight altitude of the aircrafts,

4

we restrict our analysis to the radiosonde profiles. All radiosonde and dropsonde profiles were made available for operational assimilation at weather centres (Schäfler et al., 2018). Figure 1 shows the launching position of those radiosondes that were assimilated within the IFS.
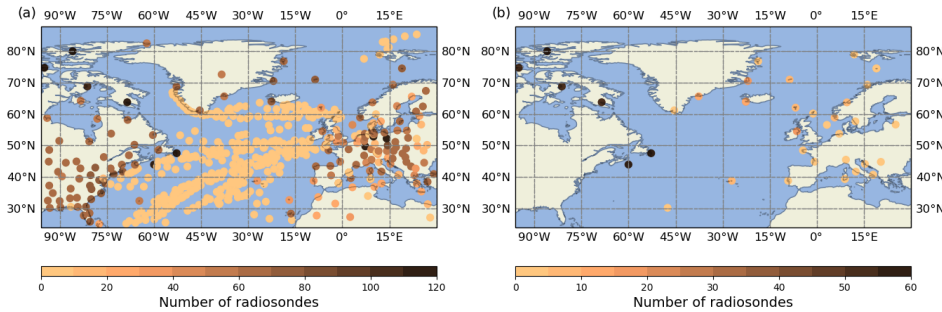

Figure 1: Positions of radiosonde launches that were assimilated by the ECMWF IFS between 17 September and 18 October 2016 for (a) all radiosondes (9729), and (b) the subset of 497 non-operational radiosondes launched during NAWDEX. The colouring denotes the number of assimilated profiles at a particular site. The scale of the colourbar changes between (a) and (b).

~~With the aim to~~To investigate the influence of the extra radiosonde observations during NAWDEX a dedicated OSE was performed with the IFS (Schindler et al. 2020). The cycled OSE covers the whole NAWDEX campaign period (17 September to 18 October) and uses IFS model cycle 43r1 (Cy43r1; ECMWF, 2016), which became operational in November 2016. The triangular–cubic–octahedral grid ($T_{Co}1279$) provides a horizontal resolution of ~9 km and 137 vertical sigma–hybrid levels that range from the surface up to ~80 km. The vertical resolution is highest in the planetary boundary layer decreases with altitude. At typical midlatitude tropopause altitudes (6–15 km; e.g., Schäfler et al., 2020) the vertical grid spacing is about 300 m. The incremental hybrid 4DVAR DA scheme used at ECMWF assimilates observations available in a 12 h time window to update a prior short range forecast in order to achieve the best possible estimation of the atmospheric's state, which is the analysis. More details about the implementation of 4DVAR in the IFS are given in Rabier et al. (2000) or in the IFS documentation (ECMWF, 2016). As in the operational ECMWF system, the B-matrix for the experiments is based on a blended combination of a climatological estimate and an estimate from an ensemble of data assimilations (EDA). The cycled OSE comprises two separate model runs. The control run (CTR) considered all routine and extra radiosondes as well as the dropsondes launched during NAWDEX. The denial (DEN) run excluded all additional observations in a region over the North Atlantic (25°–90°N; 82°W–30°E). In addition, a 25–member EDA experiment was conducted at lower horizontal resolution ($T_{Co}639$ ~18 km) for both experiments. More details on the OSE design are given in Schindler et al. (2020).

For our analysis we retrieved observation feedback files of the OSE experiment from ECMWF's observation database (ODB), which contain the (radiosonde) temperature and wind observations and their departures from the background and analysis state given as profiles using pressure ~~used~~ as the vertical coordinate. On the one hand we analyse the influence of all 9729

5

radiosondes in the operational CTR run. On the other side, the influence of the subset of 497 radiosondes in the CRL is
155 compared with the DEN experiment, where they were excluded and only passively monitored. The observation space data is
stored during the 4DVAR process at the position and time of the observations. It has to be noted that the radiosonde profiles
are not assimilated at their fully measured vertical resolution (which would be ~5 m) but at a reduced ~~limited~~ number of levels
(~50–350), which depends on the reporting type (e.g. alphanumeric, BUFR~~, highres. BUFR~~) the individual stations used for
the data transmission to the Global Telecommunications System (GTS) (Ingleby et al., 2016). About 65 % of the assimilated
160 radiosonde profiles used in this study have a low vertical resolution (<100 data points per profile) while 35 % of the profiles
exhibit up to roughly 400 levels (see Fig. S1a in Supplement). Accordingly, the distribution of the average vertical distance of
neighbouring data point in the UTLS show a bi-modal shape and varies between ~100 and 400 m in the UTLS (Fig. S1b in the
supplement). Cy43r1 does not consider the horizontal drift of the radiosondes. From the requested feedback files, we extract
those observations that are actively assimilated. Some profiles (<1 %) which do not provide temperature and wind data above
165 the 540 hPa level (~5 km) are excluded from the statistical analysis. This level is selected as it serves as a starting point for the
tropopause detection (Sect. 2.2).

2.2 Data processing and tropopause-relative coordinates

First, the observation space background (or first guess, y_{FG}) and analysis (y_{AN}) states are derived from the observations (y_O)
and departures from the first guess (dep_{FG}, referred to as innovation) and the analysis (dep_{AN}, hereinafter residuals) as follows:

170 $$\text{Innovation: } dep_{FG} = y_O - y_{FG} \qquad (1),$$
$$\text{Residual: } \quad dep_{AN} = y_O - y_{AN} \qquad (2).$$

The observation space increment is defined as the analysis minus the background state and shows ~~of~~ whether a quantity has
been increased or decreased in the DA cycle:

$$\text{Increment } = y_{AN} - y_{FG} \qquad (3).$$

175 In a next step, we derive~~d~~ the geometrical altitude from the pressure data based on the hydrostatic equation~~s~~ as described ~~given~~
in ECMWF (2016). The observation and model states are then linearly interpolated in the vertical to an equidistant 10 m grid.
The potential temperature (θ) and the squared static stability (N^2) are computed from the temperature profile using

$$N^2 = \left(-\frac{g}{\theta}\right) * \left(-\frac{d\theta}{dz}\right) \left[\frac{1}{s^2}\right] \qquad (4),$$

with $\frac{d\theta}{dz}$ the vertical gradient of θ ~~($\frac{d\theta}{dz}$)~~ in geometrical coordinates (z) and g the gravitational acceleration (~~g;~~ g = 9.81 m s^{-2}).

180 From the wind profile~~s~~, the vertical wind shear of the horizontal wind speed is~~are~~ calculated as follows:

$$\text{Wind shear} = \left(\frac{d|\overline{u}|}{dz}\right) [s^{-1}] \qquad (5),$$

with $\frac{d|\overline{u}|}{dz}$ the vertical gradient ~~$\frac{d|\overline{u}|}{dz}$~~ of the magnitude of the horizontal wind vector $\vec{u}$.

In the statistical assessment, average profiles of the parameters and increments are calculated in tropopause-relative
coordinates. Various tropopause definitions are used in the literature, which are defined based on the particular thermal,

6

185 dynamic and chemical characteristics of the UTLS (e.g., Gettelman et al., 2011). We rely on the lapse–rate tropopause (LRT),

which, by definition, points to the sharp transition of thermal stratification from the UT to the LS (e.g., Birner et al., 2002;

Tinney et al., 2022). The LRT is defined as the lowest level at which the lapse rate (i.e., the vertical temperature gradient) falls

below 2 K km^{-1}, subject to the condition that the average lapse rate from that level to any point within the overlying 2 km layer

does not exceed 2 K km^{-1} ~~(World Meteorological Organisation (WMO), 1957)~~. Th~~is~~e World Meteorological Organisation

190 ~~(~~(WMO~~)~~ definition (WMO, 1957) also comprises a further criterion to determine a secondary (or "double") tropopause,

however, in th~~e~~is presented analysis we only determine the "first" LRT. The LRT altitude is used to determine LRT~~-~~-relative

altitudes ($z_{LRT-relative}$) for each radiosonde profile, which is the difference of the geometrical height profile ($z_{geometrical}$) and

the LRT altitude (z_{LRT}) following Eq. 6:

$$z_{LRT-relative} = z_{geometrical} - z_{LRT} \quad\quad\quad\text{————(6).}$$

195 ~~In the statistical assessment averages of parameters and increments are calculated in tropopause–relative coordinates (e.g.,~~

~~Birner et al., 2002).~~ Although the LRT definition permits a robust detection of the tropopause altitude for most atmospheric

conditions, the 2 –K criterion entails some important limitations (for details see Tinney et al., 2022 and references therein).

First, the 2 –K threshold can lead to undesired false detections of the LRT altitude (in the following referred to as

"misdetection") at small temperature fluctuations which are often present in the lower troposphere (boundary layer) but also

200 occur in the mid-troposphere. To avoid LRT misdetections, the tropopause detection is performed above ~5 km (540 hPa)

altitude. Second, in situations of weak vertical temperature gradients, i.e.~~,~~. smooth transitions across the tropopause, the 2 –K

threshold is sometimes not quite met leading to LRT jumps by several kilometres for neighbouring, similar temperature profiles

(Krüger et al., 2022). This occurs typically in the vicinity of the jet streams where the tropopause altitude shows a discontinuity

or double tropopauses may occur (Pan et al., 2004; Hoffmann and Spang, 2022; Tinney et al., 2022). Hence, a slightly different

205 temperature representation in models and observations can result in large LRT altitude differences. The potential influence of

~~the~~ such misdetections is discussed in Sect. 3.2.2 and 4.

LRT altitudes are derived individually for the observations, (in the following: LRT$_{yO}$) the background (LRT$_{yFG}$) and analysis

(LRT$_{yAN}$) by following the WMO definition outlined above. Figure 2a illustrates the vertical distribution of LRT$_{yO}$ for 9729

profiles which has a bi~~-~~–modal shape in the altitude range from 6 km to 18 km with peaks at 11.5 and 15.5 km. The left mode

210 represents profiles with a high frequency (75 % of the profiles) of LRT altitudes at 10-14 km (see Fig. 2a) which is typical for

the midlatitudes in autumn (e.g., Hoffmann and Spang, 2022; Krüger et al., 2022). Its broad spectrum is related to the variability

of the midlatitude tropopause in different synoptic situations, e.g. in ridges and troughs (e.g., Hoerling et al., 1991). The right

mode (LRT > 14 km; 25 % of the profiles) with its smaller maximum indicates profiles in the subtropics. The LRT distribution

for the additional NAWDEX radiosondes (Fig. 2b) does not exhibit a corresponding second peak, due to the low number of

215 soundings conducted at latitudes < 40 °N.

~~The highest frequency, which occurs for the lower maximum, represents midlatitude profiles at 10 13 km altitude (see~~

~~colouring in Fig. 2), which are typical LRT altitudes expected from climatology in the midlatitudes in autumn (e.g., Hoffmann~~

~~and Spang, 2022; Krüger et al., 2022). The broad spectrum of the tropopause altitudes below 15 km is related to the variability~~

of the midlatitude tropopause in different of synoptic situations, e.g., in ridges and troughs. A second, smaller maximum of

220 LRT$_{yo}$ at 15–16 km is related to profiles at low latitudes in the vicinity of the subtropical jet.
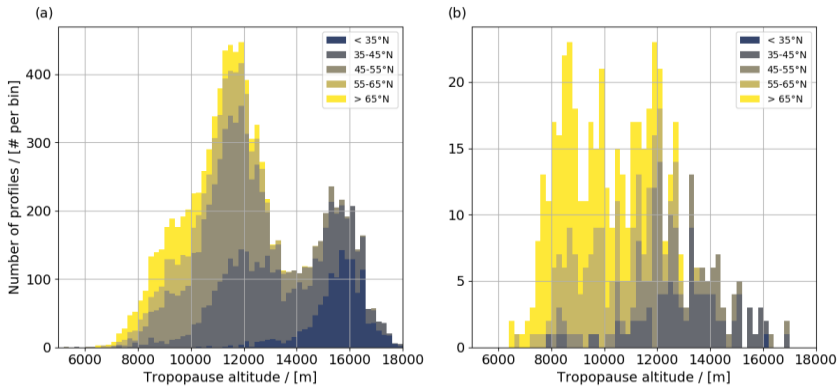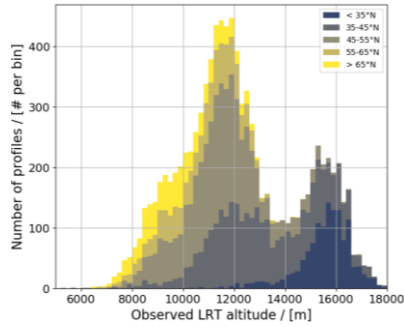

Figure 2: Stacked distribution of ~~LRT$_{yo}$~~ LRT$_{yO}$ with 0.2 km bin size for (a) all 9729 radiosondes and (b) the additional 497 radiosondes observed during NAWDEX. The colouring shows the latitude of the radiosonde stations (10° bins).

225 Figure 3 presents the mean vertical profiles of observed temperature, N², wind speed and wind shear profiles averaged in ~~LRT$_{yo}$~~ LRT$_{yO}$–relative (with respect to the observed tropopause) coordinates. These profiles outline the main characteristics of the midlatitude tropopause that are known from climatology (e.g., Birner et al., 2002; Grise et al., 2010; Hoffmann and Spang, 2022): Above a linearly decreasing temperature in the troposphere (~7 K km^{-1}), a temperature minimum of about 213 K is reached at the ~~LRT$_{yo}$~~ LRT$_{yO}$. Above, a distinct temperature inversion (0–1.5 km above LRT$_{yO}$) follows before the temperature

230 becomes roughly isothermal (up to ~5 km above the LRT) in the stratosphere. This change of stratification results in a rapid jump of N² (from 2 to 6.5x10^{-4} s^{-2}) across the LRT$_{yO}$ altitude. Wind speed continuously increases with altitude in the

8

troposphere up to a maximum (~23.5 m s^{-1}) at ~1 km below LRT$_{yO}$. Corresponding to the distribution of wind speed, the vertical shear of wind speed shear (in the following referred to as wind shear) is positive up to the wind speed maximum, then abruptly decreases beyond and reaches a distinct minimum (~5x10^{-3} s^{-1}) at about 300 m above LRT$_{yO}$. Please note that the presented data set of 9729 radiosondes provides a high data coverage (Fig. 3a, blue line) in the UTLS.
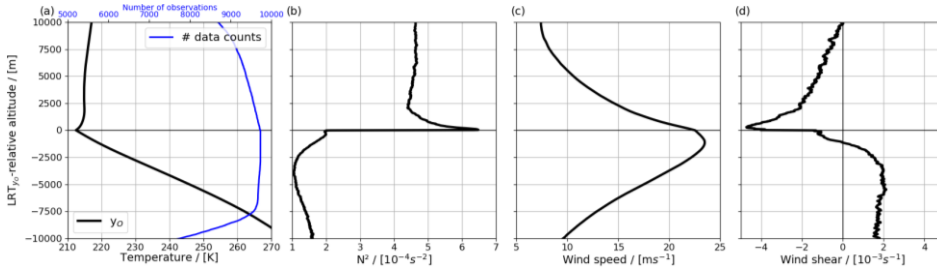

Figure 3: LRT$_{yO}$-relative mean profiles of (a) temperature (black) and number of data (blue), (b) N² and (c) wind speed and (d) wind shear using 9729 radiosondes.

A separate analysis of extratropical (LRT < 14 km) and sub-tropical (LRT > 14 km) observations reveals similar shapes for the extratropical and the overall data (see Fig. S2a-d). The sub-tropical mean profiles exhibit lower temperatures in the entire UTLS, a weaker temperature inversion in the LS and no wind maximum being located near the tropopause.

3 Results

3.1 Increments in geometrical and tropopause-relative coordinates

Figure 4 shows the time series of temperature increments over Iqaluit, Canada between 17 September and 18 October 2016, in both geometrical (Fig. 4a) and LRT$_{yO}$-relative coordinates (Fig. 4b). Iqaluit is selected as it comprises a high number of radiosonde profiles (#114, at a 6-hourly interval) and outlines the typical high tropopause altitude and wind speed variability related to the changing synoptic situations. Several strong jet stream events with wind speeds of occasionally >45 m s^{-1} passed over the station, which are accompanied by high variability of the LRT (7–13 km).
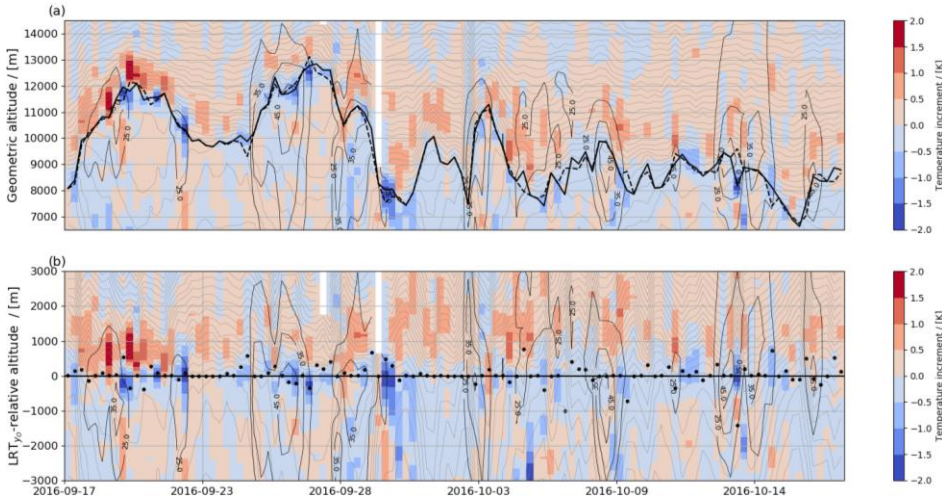
9

Figure 4: Time series (17 September –18. October 2016) of temperature increments (colour shading) at Iqaluit (63.75°N, 68.53°W, Canada) illustrated in (a) geometrical height and (b) LRT_{yO} –relative coordinates. The panels are superimposed by the observed θ (thin grey lines, $\Delta\theta$ = 4K) and wind speed (thin black contours). In (a) the black thick (dashed) line show LRT_{yO} (LRT_{yFG}) and the black dots in (b) show their difference.

In geometrical altitude strong positive (>1 K) and negative temperature increments (< -1 K) are stacked and roughly follow the tropopause. Due to the variable LRT altitude, averaging of the profiles in geometrical coordinates would blur the vertical distribution of the increments and thus, hide a potential influence on the tropopause in a statistical evaluation. However, in LRT_{yO} –relative coordinates, the negative increments can be clearly assigned to the about ±0.5 km around the tropopause, and the positive increments to 2 km thick layer above. The vertical extent of the positive and negative increments is relatively persistent along the entire time series, but the magnitude is variable. The distance between observed and background tropopause altitude is mostly in the range of ~100 m, but there are also cases with large altitude differences (>1 km, see discussion in Sect. 2.3).

3.2 Statistical assessment

3.2.1 Mean tropopause-relative influence

Figure 5 presents LRT_{yo} –relative average profiles of temperature, N^2, wind speed and vertical wind shear for the 9729 radiosondes and their model equivalents. The minimum temperature detected in a layer of ~ ±500 m around ~~LRT_{yo}~~ LRT_{yO} is

overestimated (up to 1 K) in the background profiles (Fig. 5a) confirming a warm temperature bias at the tropopause. In the LS above, the background temperature ~~in~~decreases less strongly which results in a cold model bias between 0.5–2 km above the tropopause. The weaker thermal gradients in the background are accompanied by an underrepresentation of the amplitude and sharpness of the N^2 jump across the tropopause (Fig. 5b). Wind speed is underestimated in the background throughout the UTLS (Fig. 5c), with a maximum underestimation (0.5 m s^{-1}) between –1 km and 0.5 km. The rapid decrease of wind shear above the wind maximum towards lowermost 1 km layer of the LS is less pronounced in the background. ~~The magnitude of wind shear, respectively its rapid decrease at the tropopause is much weaker in the background, leading to an underestimation of wind shear below and an overestimation above LRT$_{yO}$.~~ Figures 5–a–d show that the analysis is drawn towards the observations for all parameters at any altitude ~~of~~ in the UTLS. The slightly sharper tropopause structure reveals a positive influence of DA on the representation of the tropopause in the analysis.

Figures 5e–h shows the vertical structure of the increments. The temperature increments (Fig. 5d) imply a cooling (up to –0.25 K), between –1 km and +0.5 km around the LRT$_{yO}$, i.e., the altitude range of the warm bias. In the LS, a warming of up to 0.25 K between 0.5 and 2 km above LRT$_{yO}$ counteracts the cold bias in the model background. This impact on the temperature distribution results in negative N^2 increments (-0.15×10^{-4} s^{-2}) in a 1.5 km thick layer below the LRT$_{yO}$ and between 1 and 2 km above LRT$_{yO}$ (Fig. 5e). In the 1 km layer above LRT$_{yO}$, N^2 increments are positive with a distinct maximum of ~0.32×10^{-4} s^{-2} at ~0.5 km. Wind increments are predominantly positive in the entire UTLS (Fig. 5g), which indicates a wind speed increase in the analysis. The wind increments are stronger in the UT than in the LS peaking (0.25 m s^{-1}) at the altitude of the wind maximum and the strongest underestimation in the background (1 km layer below LRT$_{yO}$, Fig. 5c). Increments of wind shear are positive in the 2 km below, and negative in the 1 km layer above the ~~LRT$_{yO}$~~ LRT$_{yO}$ (minimum at ~500 m above LRT$_{yO}$). Between ~1.5 to 3 km above LRT$_{yO}$ are positive of comparable magnitude to the increments in the UT. A separate analysis of mid-latitude and sub-tropical increments (Fig. S3) shows that the latter are weaker. However, as the increments in both regions point in the same direction, the complete data is considered for the statistical analysis in the remainder of this article.
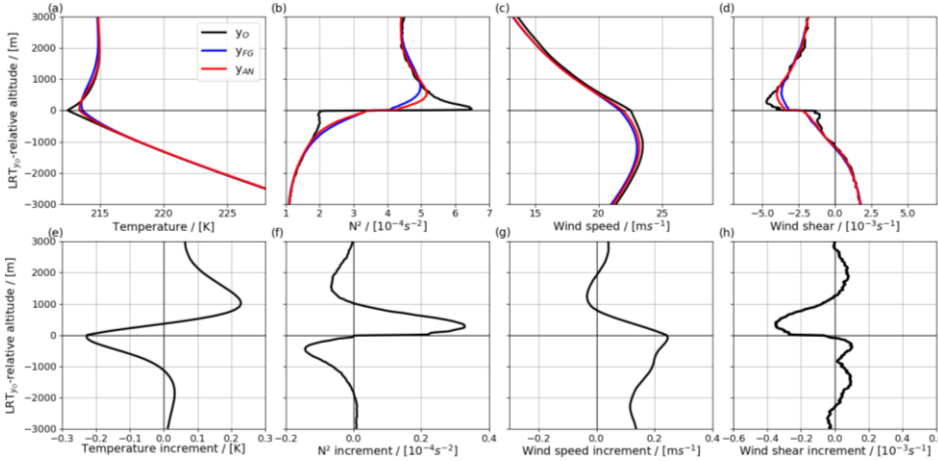
Figure 5: ~~LRTyo~~LRT$_{yO}$~~-~~relative distributions of (a) temperature, (b) N² (c) wind speed and (d) wind shear and the respective increments (e–h; Eq. 3) averaged for the 9729 profiles of observations (black), background (blue) and analysis (red).

3.2.2 Sensitivity to the LRT-~~-~~relative coordinate

Figure 5 presented profiles of the parameters and increments relative to the observed tropopause (LRT$_{yO}$). However, from Fig. 4 we have seen that the observed tropopause altitude may differ from the background (and analysis). This raises the question which LRT-~~-~~relative view is the most suitable reference for evaluating the tropopause structure. In the following, we present the distributions in different LRT-~~-~~relative reference systems and discuss their significance to evaluate the tropopause sharpness. Figure 6 illustrates the profiles of temperature and N² in the observations (Fig. 6a, e), background (Fig. 6b, f) and analysis (Fig. 6c, g) with respect to the observed LRT$_{yO}$, the background LRT$_{yFG}$ and the analysis LRT$_{yAN}$, respectively. In each case the lowest tropopause temperature as well as the strongest temperature inversion and jump in N² occurs when the "own" LRT is used. This is particularly obvious for the observed profile relative to LRT$_{yO}$ ~~(yo(LRTyo)~~; black curve in Fig. 6a, e). In addition, the background and analysis profiles have the lowest tropopause temperature and strongest inversion when viewed relative to LRT$_{yFG}$ (medium red curve in Fig. 6b) and LRT$_{yAN}$ (light blue curve in Fig. 6c), respectively. Figure 6d and h show the increments referenced to the different LRT-~~-~~relative coordinates. Each of the LRT reference systems confirms a cooling near the LRT, a warming in the LS above (Fig. 6d), and an increase of static stability just above the LRT (Fig. 6h). The differing LRT altitudes of the individual profiles in y$_O$, y$_{FG}$ and y$_{AN}$ result in small differences in the magnitude of the increments for the different LRT-~~-~~relative coordinates (discussed in further detail in Sect. 3.3). The increments are smallest when referenced to LRT$_{yFG}$.

12

As the ~~own~~ LRT$_{own}$-~~-~~relative distributions provide highest sharpness, we also consider ~~own~~ LRT$_{own}$-~~-~~relative increments (grey line in Fig. 6 d, h) to further analyze the influence on tropopause sharpness. These increments in LRT$_{own}$-~~-~~relative coordinates, which are calculated as $y_{AN}(LRT_{yAN}) - y_{FG}(LRT_{yFG})$,~~,~~ and ideally remove effects on the average increments from differing LRT$_{yFG}$ and LRT$_{yAN}$ altitudes, have a comparable structure in the LS. However, they show only a slight cooling (<0.1 K) at the tropopause and an increasing warming with decreasing altitude in the troposphere, which does not agree with increments in geometrical and LRT$_{yO}$ space (e.g., Fig. 4 and Fig. 6a). The warming in the troposphere is a systematic temperature bias that is caused by the tropopauses detected at different altitudes (either in LRT$_{yFG}$ or in LRT$_{yAN}$ or both, Fig. 6a). To emphasize the role of differing tropopause altitudes on the distribution of the increments, the four types of increments are shown for cases with similar LRT altitudes (within ±100m), which are almost identical (see overlapping dotted lines in Fig. 6d). We do not further pursue the analysis of LRT$_{own}$-~~-~~relative increments because such increments are determined after shifting the profiles with respect to the own LRT which does not correspond to real changes to the model background field in geometrical space (when the LRT$_{yFG}$ and LRT$_{yAN}$ differ). We nonetheless present this analysis ~~approach in order~~ to emphasize the sensitivity of cross-~~-~~tropopause distributions and their increments to the choice of the LRT ~~relative~~ reference and the impact of systematic LRT altitude differences. ~~As~~ LRT$_{yO}$ ~~the LRT derived from the radiosondes profiles~~ provides the most realistic representation of the tropopause altitude (Fig. 4a, Fig. 6a) and~~,~~ ~~LRT$_{yO}$~~ is used in the following to analyze the influence on tropopause sharpness in this study ~~(see Sect. 3.2)~~. In addition, the influence on the tropopause altitude is studied relative to the LRT$_{yFG}$ (Sect. 3.3).
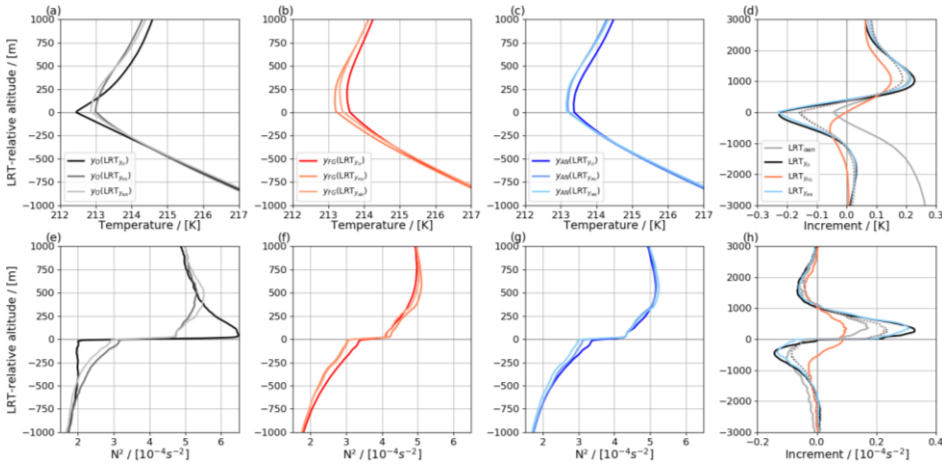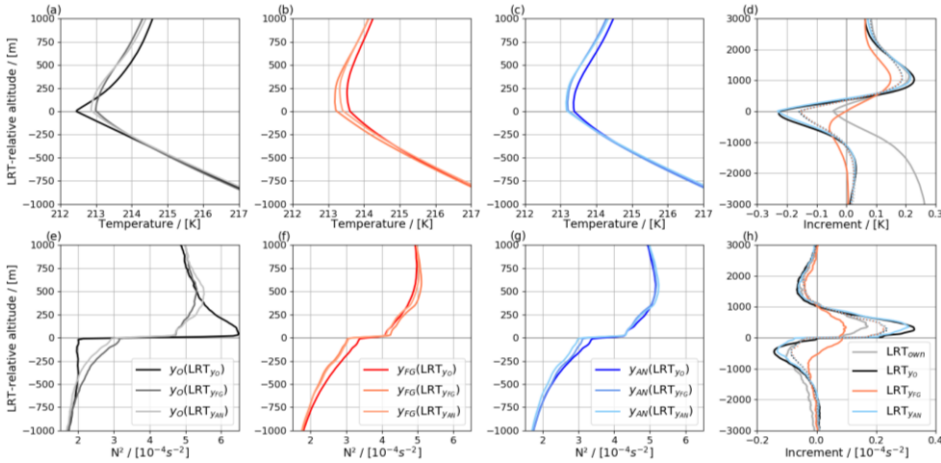
Figure 6: Mean profiles of temperature (a–c) and N^2 (e–g) for observations (a, e;), background (b, f) and the analysis (c, g) relative to LRT_{yO}, LRT_{yFG} and LRT_{yAN}, respectively (colour coded). The panels d) and h) show the associated increments. In addition, increments using the own LRTs are shown (grey, calculated as $y_{AN}(LRT_{yAN}) − y_{FG}(LRT_{yFG})$, for details see text). The dotted lines in d) and h) represent the 3712 profiles with the LRT altitudes of observations, background and analysis being within ±100 m (note that dotted lines overlap).

3.2.3 Influence on tropopause sharpness

The previous results indicated an increase of sharpness in the analysis ~~and the time series in Fig. 4~~ ~~with~~ suggested high temporal variability of the increments (Fig. 4) that is likely influenced by particular dynamical situations. Figure 7a illustrates the distribution of the observed maximum squared static stability (N^2_{max}) in the 3 km above the LRT_{yO}, which is a common indicator for tropopause sharpness (Birner et al., 2006; Pilch Kedziersky et al., 2015). N^2_{max} shows a~~n~~ uni-—modal, ~~positively skewed~~positively skewed distribution ranging from 3 to 30×10^{-4} s^{-2} with largest frequency (>200 profiles per bin) of between $6–12 \times 10^{-4}$ s^{-2} and lowest frequency (<50 profiles per bin) for 5×10^{-4} $s^{-2} < N^2_{max}$ and $N^2_{max} > 15 \times 10^{-4}$ s^{-2}. The quartiles of this distribution are used to classify the data into the smoothest ($N^2_{max,Q00–Q25}$), the intermediate ($N^2_{max,Q25–Q75}$) and the sharpest ($N^2_{max,Q75–Q100}$) tropopause cases. The observed profiles (Fig. 7b) display that the sharp class has the lowest tropopause temperature and the strongest inversion with the largest jump of N^2. On the contrary, the smoothest tropopauses exhibit a higher tropopause temperature, a weaker temperature inversion and a lower amplitude in N^2 at the ~~LRT_{yO}~~LRT_{yO}. The intermediate class depicts a comparable tropopause structure to the full data set average described in Sect. 3.2.1. In agreement with these findings the observed mean tropopause altitude for the sharp and smooth classes are 12750 m and 11580 m, respectively, which suggests that the sharp (smooth) tropopauses can be related to ridge (trough) situations characterized by high (low) tropopause altitudes.
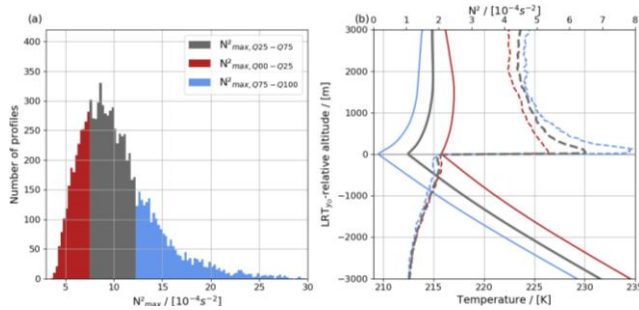
14

Figure 7: (a) Distribution of N^2_{max} as observed in the 3 km layer above LRT_{yO} with bin size 0.5×10^{-4} s^{-2}. Quartiles of the N^2_{max} distribution in (a) are used to classify tropopause sharpness: smoothest (red; Q00–Q25), intermediate (grey; Q25–Q75) and strong sharpness (blue; Q75–Q100). **(b)** the corresponding ~~LRT_{yO}~~LRT_{yO}— relative mean profiles of observed temperature (solid) and N^2 (dashed).

For each class of N^2_{max}, the mean vertical profile of innovation (Eq. 1), increment (Eq. 3) and residual (Eq. 2) for temperature and N^2 relative to LRT_{yO} are presented in Fig. 8. We first focus on the intermediate tropopause sharpness. In the UT, the temperature innovations are weak, negative and vertically nearly constant (Fig. 8a, about –0.1 K) before they reach a minimum of about –1.2 K at LRT_{yO} indicating a warm bias at the background tropopause. Above, the innovations strongly increase and become positive at ~0.5 km above the LRT_{yO} before a maximum cold bias of ~0.3 K is reached at 0.8 km altitude. The temperature increments (Fig. 8b) correspond to the findings in Fig. 5 with the negative increments around the LRT_{yO} counteracting the warm bias, and the positive increments above decreasing the cold bias. In the ±0.5 km around LRT_{yO} large N^2 innovations between –2 to 3×10^{-4} s^{-2} illustrate the strong underestimation of tropopause sharpness in the background (Fig. 8d). The average positive (above LRT_{yO}) and negative (below LRT_{yO}) N^2 increments (Fig. 8e) for the intermediate profiles agree in shape and magnitude to the structure of N^2 increments given in Fig. 5. Apparently, they lead to a sharpening of the tropopause. Increments are much smaller than the innovations (~20 % for temperature and 10 % for N^2) which explains that the vertical structure of the innovation is preserved in the residuals (Fig. 8c, f). For the smooth and sharp classes (blue and red lines in Fig. 8), innovations, increments and residual have a similar vertical distribution but show weaker, respectively stronger amplitudes. For instance, temperature increments are about –0.3 K (–0.1 K) at the tropopause for the sharp (smooth) class and about 0.3 K (0.1 K) for the maximum above, in the LS. The influence is stronger where the background biases are strongest.

15
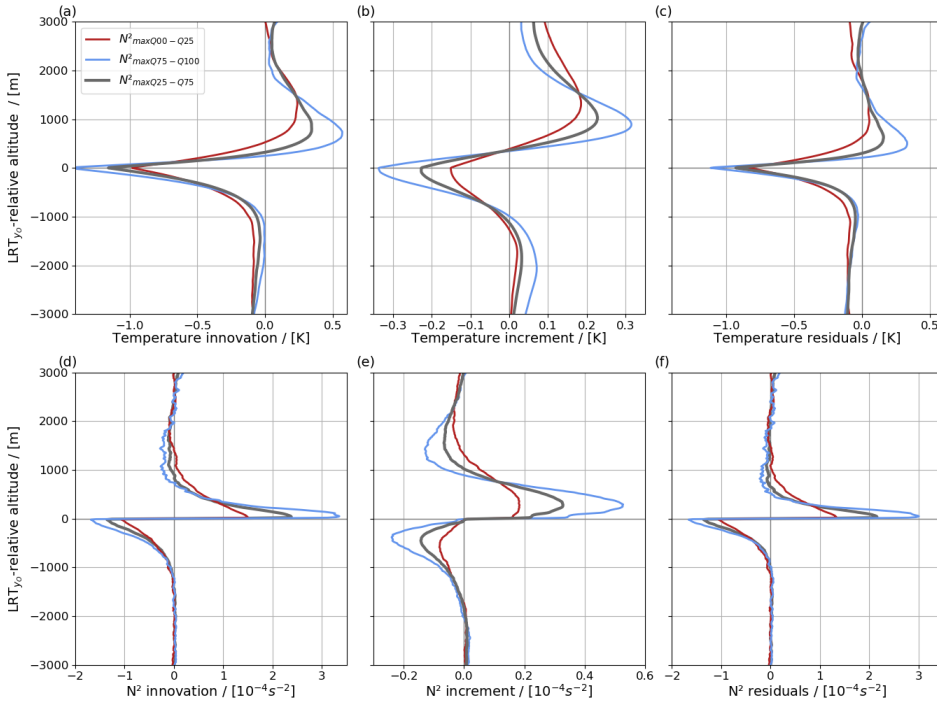
Figure 8: LRT$_{yO}$ -relative mean profiles of innovations, increments, and residuals for (a–c) temperature and (d–f) N^2 for the classes of N$^2_{max}$ defined in Fig. 7.

Figure 9 ~~investigates~~ illustrates the variability of wind speed in the radiosonde data set, ~~which varies strongly at the individual~~ ~~stations over time (Fig. 4)~~. Average wind speeds in a layer of ±3 km around LRT$_{yO}$ range from nearly 0 to 60 m s^{-1} with the highest frequency between 5–25 m s^{-1} (Fig. 9a). Quartiles of layer- mean wind speed divide the data set into weak (wind$_{Q00-Q25}$), intermediate (wind$_{Q25-Q75}$) and strong (wind$_{Q75-Q100}$) winds. ~~Consequently, t~~The weak wind class shows vertically fairly constant low wind speeds (<10 m s^{-1}). While the intermediate ~~wind~~ class exhibits a comparable shape of the mean wind profile as ~~structure to~~ the full data set (Fig. 5), the strongest wind class depicts a pronounced wind maximum (>40 m s^{-1}) at –1 km altitude below the LRT$_{yO}$ expressing strong jet stream winds. While wind shear in the weak wind class is small, the intermediate and strong wind class have a decrease of vertical wind shear from positive values below the wind maximum to negative values above, with a peak in the LS. All classes show an increased reduction of wind speed above the LRT which is associated with a step change in vertical wind shear. The mean ~~wind~~ profiles of innovation, increment and residual for each class are shown

16

in Fig. 10. The positive wind innovations of all classes across the UTLS express the underestimated wind speeds in the

385 background (Fig.10a, see also Fig. 5). Innovations in the UT are generally larger than above in the LS and peak at the

tropopause. Maximum innovations range between 0.5 m s^{-1} for the weak and 1.2 m s^{-1} for the strong wind class. The

predominately positive wind speed increments throughout the UTLS (Fig. 10b) represent a wind increase in the analysis which

is largest in the 500 m layer below LRT$_{yo}$ LRT$_{yO}$ ranging between 0.1 m s^{-1} for the weak and 0.45 m s^{-1} for the strongest wind

class. The positive wind speed residuals (Fig. 10c), show that a slow wind speed bias remains in the analysis, however the

390 weaker residuals than the innovations observed for each class point to an improvement of wind. For the strongest winds the

innovations are reduced by roughlyup to 50 40 %. In the layer with increased negative shear directly above the LRT, there are

particularly strong negative innovations for the class of the strongest winds, which are associated with increased negative shear

increments. After the assimilation, wind shear residuals show little variability across all classes.
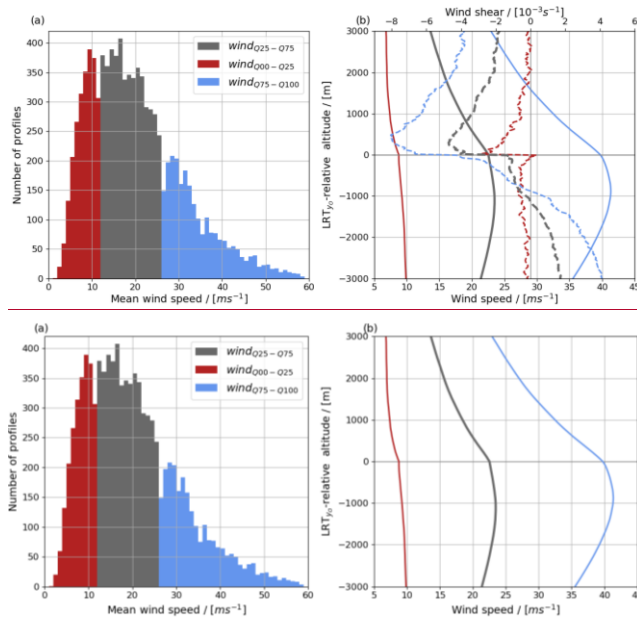

Figure 9: (a) Distribution of mean observed wind speed in the ±3 km above and below LRT$_{yO}$ with 1 m s^{-1} bin size. Quartiles of the

distribution in (a) are used to distinguish wind classes: weakest wind (red; Q00–Q25), intermediate wind (grey; Q25–Q75) and strongest

wind (blue; Q75–Q100). (b) The corresponding LRT$_{yO}$-relative mean profiles of observed wind speed (solid) and wind shear (dashed) per
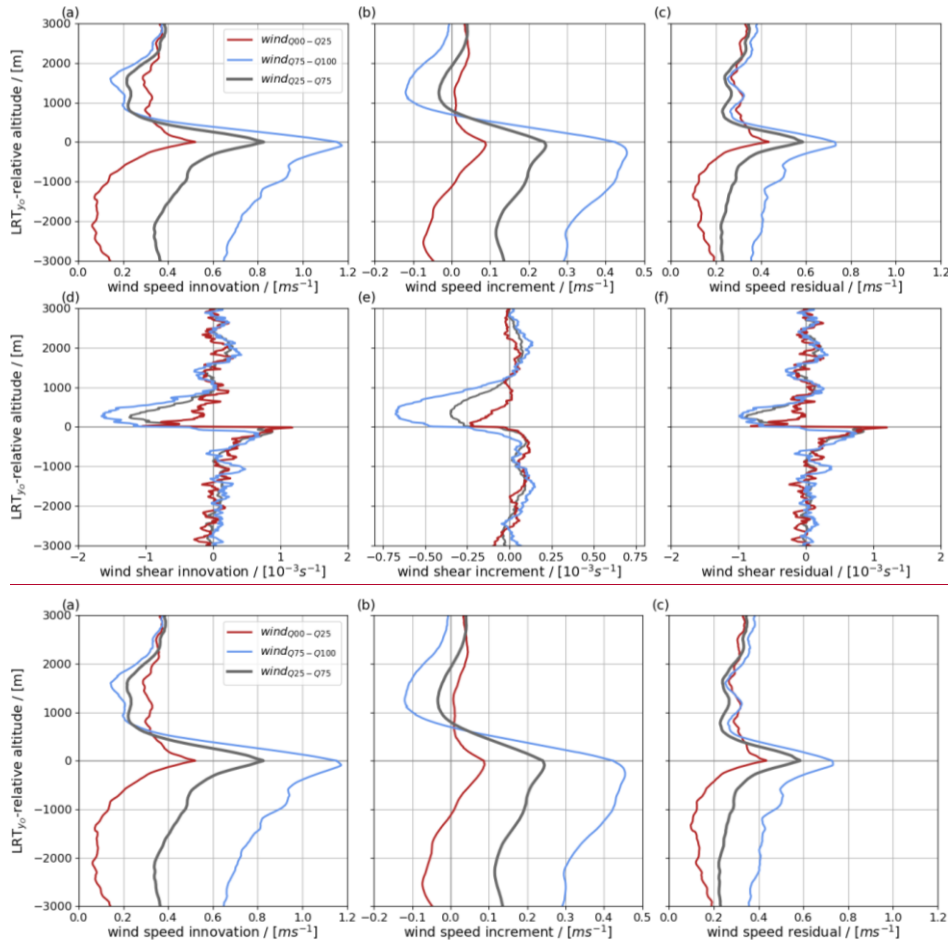
400 class.

Figure 10: LRT$_{yO}$-relative mean profiles of wind speed and wind shear (a,d) innovations, (b,e) increments and (c,f) residuals per classes defined in Fig. 9.

405   **3.3 Influence on the tropopause altitude**

In the following, we investigate how the modification of the temperature profile (Fig. 5e) affects the tropopause altitude. ~~In the subsequent~~Hereafter, LRT altitude differences between the observations and the background (LRT$_{yO}$−LRT$_{yFG}$) are referred

18

to as "LRT innovations" according to Eq. (1). LRT altitude differences between the observations and the analyses are referred to as "LRT residuals" (LRT_{yO} – LRT_{yAN}; Eq. 2), respectively. An overview of LRT innovations for the entire data set is given

410 in Fig. 11a providing a symmetric normal distribution centered near zero (–26 m). For 43 % of the profiles, the LRT innovations are in range of ±100 m. For about 10 % of the profiles, LRT differences are larger than 1 km. In order to prevent ~~the~~ an impact of mis-detected LRTs ~~misdetections~~ (see discussion in Sect. 3.2.2), which are most likely for unusually large LRT differences, the following evaluation is restricted to 8778 profiles (~90 %) which provide LRT altitude innovations within ±1 km.

415 First, we compare the LRT innovations (Fig. 11b, subset of Fig. 11a) and LRT residuals (Fig. 11c)~~: In Fig. 11b (subset,~~ which are colour-coded ~~of Fig. 11a)~~ at different intervals of LRT innovations ~~are colour-coded~~. The grey interval reflects LRT innovations within ±100 m, while bluish colours represent profiles where the observed LRT is higher than the background and reddish colours vice versa. ~~The~~ Comparing the colour distributions ~~ing of~~ Fig. 11b and ~~is preserved in~~ Fig. 11c allows ~~to~~ interval changes to be identif~~ied~~y ~~whether profiles have changed the interval, and~~. ~~O~~only a small fraction of profiles change

420 the intervals~~ exhibits an LRT_{yAN} vertically closer or further away from the LRT_{yo} with respect to the classes of LRT innovations~~. The distribution of LRT innovations shows a clear maximum near zero with a frequency corresponding to about 3000 profiles per bin, and the frequency decreases towards the edges of the distribution to ~10 profiles per bin. The distribution of LRT residuals (Fig. 11c) shows a slightly increased number (+15 %) of profiles within ±100 m indicating an improved tropopause altitude in the analysis and this improvement~~. This~~ is ~~also~~ confirmed by a slightly narrower shape of the gaussian

425 fit of the LRT residuals compared to the LRT innovations.

For the different intervals of LRT innovations in Fig. 11, Table 1 provides the number of profiles, the mean LRT_{yO} altitude, and the mean innovation, residual and improvement~~s~~. Except for the interval with the smallest innovation (±100 m; grey interval), the average innovation is larger than the residual which implies a vertical shift of LRT_{yFG} towards LRT_{yO}. The generally positive influence is supported by the depicted improvements, defined as the absolute difference of innovation and

430 residual. Interestingly, both, the LRT altitude shift and also the improvement grow with increasing distance between LRT_{yO} and LRT_{yFG}. ~~For the interval of smallest innovations, Table 1 shows a slightly higher average LRT residual compared to the innovation, which is a result of a small number of profiles with deteriorated tropopause altitude.~~
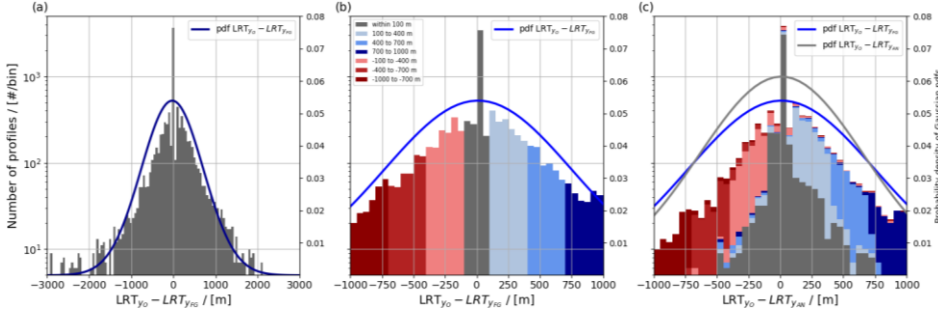
19

Figure 11: Distribution of (a) LRT innovations in the full data set, (b) as in (a) but for the range ±1000 m and (c) LRT residuals. The colour coding reflects intervals of LRT innovations shown in (b) and is reused in (c) to visualize the LRT altitude change in LRT_{yAN} (for details see text). Gaussian probability density function (pdf) are given in lines of dark blue, blue and grey, respectively, for 50 m bins. Note the log-scale of the y-axis.

LRT innovation intervals [m]	Number [#]	LRT_{yo} [m]	Innovation [m]	Residual [m]	Improvement [m]
All profiles	9729	12226	–26	12	79
–1000 to –700	209	12806	–812	–420	371
– 700 to –400	473	12425	–525	–277	227
– 400 to –100	1309	11963	–237	–121	76
– 100 to 100	4196	12064	2	18	–33
+ 100 to 400	1637	11706	222	175	36
+ 400 to 700	627	12456	517	349	153
+ 700 to 1000	227	13607	830	555	271

Table 1: Number of profiles, averaged observed LRT altitude, innovation, residual and improvement for different intervals of LRT innovation (see Fig. 11 b). The improvement is defined as the averaged $|LRT_{yo}-LRT_{yFG}|-|LRT_{yo}-LRT_{yAN}|$, so positive values reflects an improved LRT altitude in the analysis.

In order to understand how the LRT altitude changes are related to the changes of the background temperature profile, the average temperature increments for the individual intervals of LRT innovations are presented with respect to the LRT_{yFG}-relative altitude (Fig. 12). For small LRT innovations (within ±100 m, grey line in Fig. 12 a,b), the temperature increments are negative at the LRT (–0.15 K) and positive in the LS (0.2 K) (analogous to the sharpening influence discussed in Sect. 3.1 and Sect. 3.2) which does not lead to major changes in the LRT altitude in this interval (see Tab. 1). With increasing LRT innovations, the altitudes of the peaks in the LRT_{yFG}-relative increments are vertically shifted (compare coloured profiles in

20

~~Fig. 12).~~ In case of the negative LRT innovations (Fig. 12a), which means that the observed LRT_{yO} is located lower than the background LRT_{yFG}, we observe positive increments (warming, 0.3–0.6 K) at and above LRT_{yFG} and negative increments (cooling, 0.2–0.4 K) below LRT_{yFG}. As the strongest negative increments agree with ~~are located near~~ the observed tropopause

450   altitude (dotted lines in Fig. 12a), peaks in the increments are shifted downwards and show slightly higher maxima for increasing~~more~~ negative LRT innovations (red profiles in Fig. 12a). In contrast, positive LRT innovations (i.e., LRT_{yO} located above LRT_{yFG}; Fig. 12b) exhibit negative increments (–0.3 to –0.4 K) above LRT_{yFG}, and positive increments below the LRT_{yFG}. Here, the increment peaks are shifted upwards for more positive LRT innovations (blue profiles in Fig. 12b) in agreement. ~~The peaks of the negative increments coincide~~ with the altitude of the observed tropopause. ~~This confirms the~~

455   ~~indication of Fig. 4 that the negative (positive) increments are mostly aligned to (located above) the observed tropopause altitude. In summary, we found a positive influence of DA on the tropopause altitude, which is expressed by the smaller LRT residuals than LRT innovations. The potential relationship between this LRT altitude improvement and the vertically shifted temperature increments is discussed in more detail in Sect. 4.~~
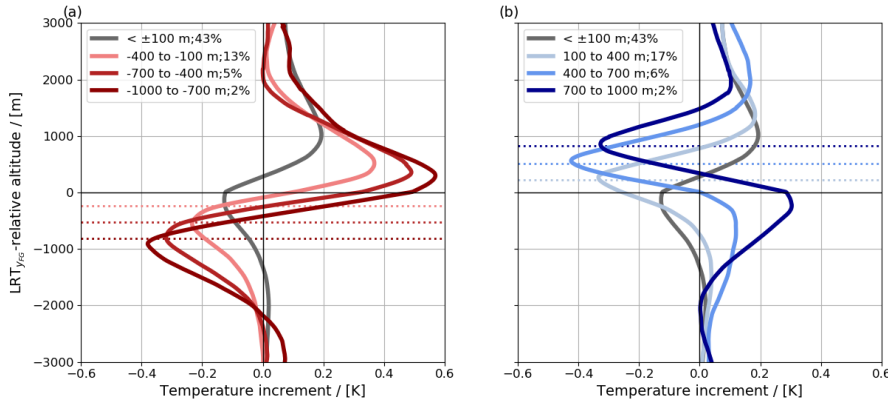

460   **Figure 12:** Mean temperature increments with respect to LRT_{yFG}-relative altitude for intervals of LRT innovations (colour coding, following Fig. 11). (a) Negative LRT innovations ($LRT_{yo} < LRT_{yFG}$, in red) and (b) positive LRT innovations ($LRT_{yo} > LRT_{yFG}$, in blue). The grey lines show the increment for LRT innovations within ±100 m. In (a) and (b) the averaged LRT_{yFG}-relative altitude of LRT_{yo} for each interval is depicted by the dotted lines.

3.4 Attributing the influence to the radiosondes

465   The presented results revealed that DA sharpens the tropopause at the location of the radiosondes, which provides a strong indication that this influence is related to the information contained in the radiosondes. However, a potential contribution of other observations cannot be excluded. For this reason, we compare the profiles and increments of the 497 NAWDEX radiosondes in the CTR run with the DEN run, in which they are denied and only passively monitored. The average profiles

21

of observed temperature and N² of the 497 NAWDEX profiles (Fig. 13) are comparable to the average profiles of the 9729

470 radiosondes (Fig. 5) with a similar magnitude of the ~~N²~~ jump of N² and an alike decrease of wind shear across the tropopause. The average minimum temperature at the tropopause (Fig. 13a) is slightly higher (by ~2 K) which is related the observation locations of the NAWDEX radiosondes at higher latitudes (Fig. 2a and b) where the tropopause is typically lower and warmer. The additional radiosondes show a pronounced wind maximum below ~~LRT$_{yo}$~~ LRT$_{yO}$ of about 29 m s^{-1} (Fig. 13c), which compared to the lower wind speeds in the complete data set (Fig. 5c), indicates the occasionally strong jet streams in the focus

475 of the NAWDEX campaign (Schäfler et al., 2018).

The increments for temperature, N², wind speed and wind shear for the subset of 497 additional NAWDEX radiosondes are presented in Figures 13e-f. The CTR run exhibits a vertical structure that is comparable to the complete data set as discussed in Sect. 3.2: Temperature increments are negative (–0.25 K) around the observed tropopause, and positive in the LS (1–2 km above LRT$_{yO}$). Accordingly, the CTR N² increments possess a similar distribution with a 1 km layer of positive increments

480 just above LRT$_{yO}$ with a maximum of 0.3×10^{-4} s^{-2}, weak negative increments (-0.2×10^{-4} s^{-2}) in a 1 km layer beneath LRT$_{yO}$ and increments around $\pm 0.1 \times 10^{-4}$ s^{-2} beyond the 1 km layers. The vertical structure of CTR wind speed increments for the NAWDEX radiosondes also agree with the complete data set ~~(c.f. increments for the strongest wind class, Fig. 10b)~~, with a positive increment (~0.2 m s^{-1}) in the UT and a negative increment in the LS. Wind shear increments in the CTR run are also similar comparable to the full data set, but the ~~minimum~~ wind shear increments just above the LRT$_{yO}$ show~~s~~ slightly lower

485 values (-0.5×10^{-3} s^{-1}) compared to Fig. 5h. The increments of temperature, N², wind speed and shear in the DEN run (Fig. 13) are weaker at each altitude ~~and~~ but tend to pick up a similar vertical distribution. This implies that the main contribution of the tropopause sharpening and influence on wind comes from the assimilated radiosondes, but the non-zero DEN increments indicate other observations also ~~to~~ influence the tropopause structure in the same direction. This may be due ~~to~~ either to the remote impact of operational radiosondes or to dropsonde observations of which a larger number were deployed during

490 NAWDEX (see Schindler et al., 2020). Further contributions of assimilated aircraft observations and GPS radio occultation data are also conceivable. ~~other vertically resolved observations as e.g., GPS radio occultation or dropsonde observations.~~
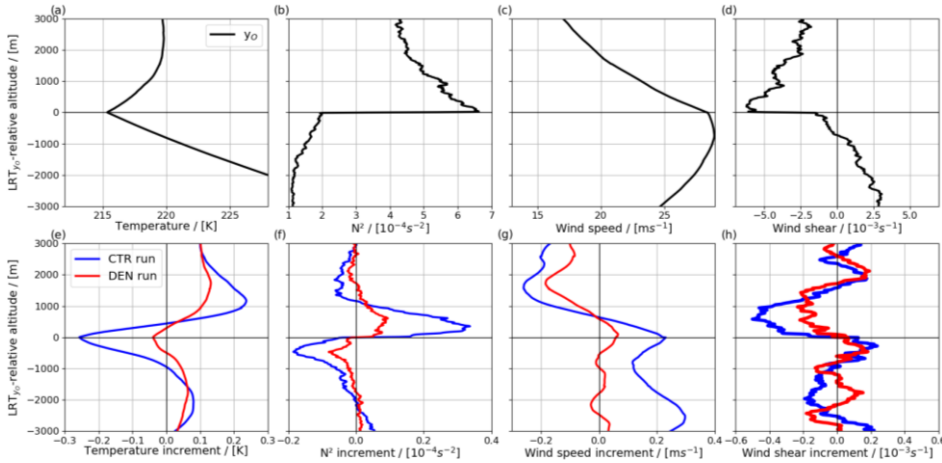
22

Figure 13: LRT$_{y_O}$-relative mean profiles of the (a) temperature, (b) N², (c) wind speed and (d) wind shear as observed by the 497 NAWDEX radiosondes and the respective increments of (e) temperature, (f) N² and (g) wind speed and (h) wind shear for the CTR (blue) and DEN (red) experiment.

4 Discussion

In this study we evaluate the influence of DA on the structure of the tropopause in the ECMWF IFS based on 9729 midlatitude radiosonde profiles. The statistical evaluation of observed temperature and wind as well as derived N² and wind shear in (thermal) tropopause-relative coordinates reproduces the typical sharp vertical gradients at the midlatitude tropopause (Birner et al., 2002). The LRT altitudes between 6–18 km during fall are considered to be representative for this area and season (e.g., Krüger et al., 2022).

To address the influence of DA on tropopause sharpness and altitude, the radiosonde and model states are transferred to tropopause-relative coordinates. The selection of a suitable reference is challenging because the tropopause-relative distributions in observations, background and analysis vary in the different LRT-relative coordinates due to slightly varying individual tropopause altitudes. Such LRT altitude differences may result from either mis-detections caused by slight temperature fluctuations in the upper troposphere or from differences in the 3D temperature distribution, e.g., in the vicinity of the jet streams. Since the origin of these LRT altitude differences cannot clearly be identified and they affect the evaluation of the tropopause altitude influence, we only consider LRT differences in the range of ±1 km. The sharpest tropopause in the observations, background and the analysis occur when viewed with respect to the "own" LRT. However, it turned out that the impact on sharpness cannot be assessed in own LRT coordinates as LRT altitude deviations in the background and analysis

23

profiles cause a spurious tropospheric temperature bias. The observed tropopause as a reference provides the most accurate representation of the tropopause to evaluate the influence on tropopause sharpness. However, temperature increments with respect to the background tropopause are useful to better understand changes to the background profile. This shows that the tropopause reference system needs to be purposefully and carefully selected and uncertainties related to the tropopause

515 detection need to be considered.

In this study, we highlight that DA improves the underestimated sharpness of the tropopause by introducing systematic changes to the background temperature and wind profile. The temperature innovations indicate a warm bias at the tropopause and a cold bias in the LS. In this layer of the sharp reversal in thermal stratification, N^2 is overestimated (underestimated) by the background below (above) the tropopause confirming findings by Birner et al. (2002). The temperature increments tend to

520 move the background temperature towards the observations by decreasing the tropopause temperature by ~0.25 K and increasing it with a similar magnitude in the LS above, which was already indicated by Radnóti et al. (2010) for the UTLS. The accompanied increase of N^2 (0.3×10^{-4} s^{-2}) in a 1 km layer above the tropopause and a decrease of N^2 (~0.2×10^{-4} s^{-2}) in the uppermost troposphere is equivalent to a tropopause sharpening, which is consistent in shape and magnitude with Pilch Kedziersky et al. (2016). Although the increments clearly reduce the background biases, the influence is rather small (~10 %)

525 compared to the innovations. The remaining LS cold bias in the analysis (0.2 K) corresponds to previous assessments (Radnóti et al., 2010) and is driven by radiative cooling due to water vapor (Sheperd et al., 2018; Bland et al., 2021), which is systematically overestimated at those levels (Krüger et al., 2022). Recent changes at the ECMWF reduced but not fully removed the bias in the IFS (Polichtchouk et al., 2021). The remaining LS cold bias in the analysis (0.2 K) corresponds to previous assessments (Radnoti et al., 2010; Shepherd et al., 2018; Bland et al., 2021). The warm bias (1 K) at the tropopause

530 in the IFS was related to the finite vertical resolution of the IFS incapable of fully resolving the tropopause (Ingleby et al., 2016), the assimilation of warm-biased aircraft data at tropopause flight levels (Ingleby et al., 2017) and the moist bias in the LS of the IFS (Bland et al., 2021). The magnitude of the warm bias (about 1.2 K) at the tropopause is about 2-3 times stronger than the corresponding warm bias reported in Bland et al. (2021). The warm bias at the tropopause (1.2 K) is in line with Ingleby et al. (2016). However, the magnitude of the warm bias is roughly 2-3 times smaller compared to Bland et al. (2021).

535 This difference may be related to vertical smoothing of the radiosonde profiles in Bland et al. (2021), which could lead to a higher tropopause temperature (König et al., 2019). Large N^2 biases (-2 to 3×10^{-4} s^{-2}) in the analysis are found in the ± 0.5 km layer around the tropopause. In addition, we show that the magnitude of the tropopause sharpening depends on the dynamic situation. For sharper tropopauses, which are typically related to higher and thus colder tropopauses occurring in situation of upper-level ridges (Hoerling et al., 1991; Pilch Kedziersky et al., 2015), temperature (and thus N^2) increments, innovations

540 and residuals are larger.

Positive wind innovations (~about 1 m s^{-1} near the tropopause) reveal the existence of a slow wind bias in the background, particularly for the wind maximum, which confirms findings by Schäfler et al. (2020) and Lavers et al. (2023). The observed vertical wind shear profile is characterized by positive values below and negative above the wind maximum as well as by a sharp increase of negative shear across the tropopause. The enhanced (negative) shear in the 1 km layer above the tropopause

24

545 in the observations is also present in the ECMWF, which is consistent with previous findings (Schäfler et al., 2020; Kaluza et al., 2021); its magnitude, however, is considerably weaker in the background and analysis as compared to the observations.~~In the layer of ± 1 km around the tropopause the observations show a sharp contrast of vertical wind shear from positive shear in the up to the wind maximum in the UT to negative shear in the LS, with the strongest gradient of shear located at the tropopause.~~ We find positive wind speed increments in the UT with a peak at the tropopause (0.2 m s^{-1}) leading to a corresponding

550 acceleration of wind speed and nearly unchanged winds in the LS. ~~Wind shear is underestimated by the background in the uppermost troposphere and overestimated in the lowermost stratosphere.~~ The wind shear increments are ~~associated by~~ positive ~~shear increments~~ just below the tropopause and negative ~~shear increments~~ in a 1 km layer above the tropopause. The generally positive influence of DA on the wind profile at all altitudes ~~on the wind~~ is depicted by ~~the~~ smaller ~~analysis~~ residuals than ~~background bias~~innovations for all ~~different~~ wind speed ~~classes~~. However, we find that high wind speed situations are

555 characterized by an increased wind bias in the background around the tropopause (underestimation of 1.2 m s^{-1}) which is reduced by ~~almost~~ about 45~~0~~ % in the analysis. This confirms Schäfler et al. (2020) who speculated that large wind errors near the jet stream in IFS short-range forecasts are reduced in the analysis. The stronger positive impact on wind for high wind speed situations was recently demonstrated by Lavers et al. (2023).

In a further investigation, we found that the influence on the temperature profiles also affects the vertical position of the

560 tropopause altitude in the analysis. While for individual profiles the LRT altitude difference of observations, background and analysis can exceed 1 km, the average differences are small (< 50 m) compared to the vertical resolution of the model of about 300 m at the tropopause. Bland et al. (2021) showed a higher tropopause altitude of about 200 m in IFS analyses using a previous model cycle (Cy41r2) and 3204 radiosondes that are a subset of the data set analysed in this study. Again, the differing results may be related to the vertical smoothing of the radiosonde profiles (König et al., 2019). In this study, profiles were

565 interpolated to a 10 m vertical grid to guarantee an accurate detection of the LRT. Certainly, the comparison is affected by representativeness errors when comparing point measurements to grid-average NWP values as discussed (Weissmann et al., 2005; Hodyss and Nichols, 2015; Janjic et al., 2018). Such an effect could be partially addressed through vertical averaging of the profiles, however, the vertical resolution of the assimilated data (100-400m, see Fig. S1) is already close to the model grid spacing in the UTLS (~300 m).

570 We reveal a positive influence of DA on the representation of the LRT altitude in the analysis that is closer to the observations than the background (or first-guess). ~~In case the LRT altitude of background and observation is comparable (in the range of ±100 m) the analysis tropopause altitude shows a slightly higher distance to the observed tropopause (16 m), which is small compared to the vertical resolution of the IFS.~~ In case of increased tropopause differences (LRT innovations >100 m), the analysis shows a systematically improved LRT altitude whereas the improvement grows with increasing LRT innovation. The

575 vertical shift of the temperature increments with respect to the background tropopause agrees with the resulting LRT altitude changes in the analysis: If the observed tropopause lies below the background tropopause, the region below is cooled, which leads to a lower LRT in the analysis. In contrast, if the observed LRT is located above the background, the region above is cooled, which, on average, shifts the analysis LRT upwards. Bland et al. (2021) and Schmidt et al. (2010) show that local

25

temperature changes in the UTLS affect the tropopause altitude. For instance, a cooling of the LS and a warming of the UT
leads to a higher tropopause altitude in models. The opposite effect, i.e. a lower located tropopause, is true in case of a cooling
of the UT and a warming of the LS. The changes of the temperature observed in this study that are induced by the DA thus
provide a reasonable explanation for the changed representation of the tropopause altitude.

The analysis of a subset of 497 NAWDEX profiles considered in a data denial OSE allowed the sharpening to be attributed
directly to the assimilation of the radiosondes. The control run increments which assimilated NAWDEX radiosondes, showed
a similar shape and magnitude as the full data set. The increments in the denial run, where the non-operational radiosondes
were only passively monitored, are much weaker, but the positive and negative increments pointing in the same direction as
the control run. Hence, the radiosonde assimilation provides the major contribution to the increments (and thus the sharpening
influence), which likely holds for the entire data set of the presented results. The non-zero increments in the denial run might
be related to the assimilation of other observations, for instance GPS–RO data (Pilch Kedziersky et al., 2016), or to the
contribution of the routine~~ly~~ radiosondes and aircraft data that are assimilated in the same assimilation time-window at a
near~~close~~-by location. A more sophisticated OSE with more observations and different observation types to be denied would
be required for a deeper investigation of this effect. The approach of assessing an OSE in observation space allows to evaluate
the influence of the observations on temperature and wind distributions on a local scale. However, the $\boldsymbol{B}$-matrix in hybrid
4DVAR schemes spreads information of assimilated observations also horizontally ~~and vertically~~ in space and time. This poses
the question to which extent the sharpening influence on the temperature and wind gradients in the UTLS, but also on the
tropopause, affects not only locally but also the surrounding region in the model. To answer this question, the authors work on
an evaluation in model space in a subsequent study.

5 Conclusion

Weather and climate predictions ~~strongly~~ rely on an accurate representation of the sharp cross-tropopause gradients of
temperature and wind. However, the initial conditions of current NWP models substantially underestimate these gradients, i.e.,
the sharpness of the tropopause. DA is known to correct for erroneous vertical distributions of temperature and wind in the
model background forecast. In this study, we address the question whether DA (positively) influences the sharpness and
altitude of the midlatitude tropopause. For this purpose, a large data set of radiosonde observations observed during a one-
month period in fall 2016 is compared with ECMWF IFS background and analysis profiles. The main conclusions of this study
following the research questions raised in the Introduction are summarized below:

1. How is tropopause sharpness represented in background forecasts and what is the influence of DA on the analysis? Does
the diagnosed temperature and wind influence depend on the tropopause structure and vary in different dynamic situations?
The tropopause-relative analysis of the DA influence on temperature, N^2, wind speed and shear using the 9729 radiosondes
shows that the tropopause is sharpened. This sharpening is described by an average cooling at the tropopause (0.25 K) and a

26

heating (0.25 K) of the LS (0.5 to 1.5 km above the observed tropopause). These increments ~~are~~ corresponding to an increase of N^2 (0.3×10^{-4} s^{-2}) in a 1 km layer just above the tropopause. We furthermore find an acceleration of wind speed (~0.2 m s^{-1}) which is most pronounced at the altitude of the highest observed wind speeds. The sharp contrast of wind shear from positive values ~~in the UT to~~below and negative values ~~at the LRT~~ above the wind maximum and especially in the lowermost LS is increased. For each parameter, the increments sharpen the tropopause, however, the influence is found to be small compared to the magnitude of the model background biases. We further uncover a sensitivity of the influence to different dynamic situations. Larger increments, but also larger innovations/residuals, are connected to sharper (N^2_{max} used as indicator) tropopauses, that are ~~typically~~ associated with ridge situations (high tropopause), while a weaker influence is observed for smoother classified tropopauses, which are related to troughs. ~~In addition, we investigated the~~The influence on the cross-- ~~tropopause wind distribution is characterized by a~~, ~~and detect~~ reduction of the slow wind bias across the tropopause. The largest positive influence occurs fFor strong jet stream wind situations, with a, ~~we find the largest influence corresponding to a~~ reduction of both, the slow wind bias ~~by ~50 %~~ in the UT by about 40 % and, of a similar magnitude, the wind shear error in the 1 km layer above the LRT.

2. Does the influence on the temperature profile affect the tropopause altitude?

A unique aspect of this study is that the DA influence on the tropopause altitude is systematically addressed. On average, tropopause altitudes differences of observations to the background and the analysis are within 50 m and for about 90 % of the profiles the tropopause altitudes agree within ±1 km. We find a positive influence of DA on the tropopause altitude in the analysis, which is expressed by a narrower distribution of the LRT residuals compared to the LRT innovations. With increasing difference between observed and background tropopause, we detect a stronger positive influence on the tropopause altitude in the analysis. The altitude improvement can be attributed to systematic temperature increments relative to the background tropopause which cause a distinct vertical shift depending on the position of the observed tropopause. If the background tropopause is located either higher or lower than the observed, the temperature increments pull the background towards the observed tropopause.

3. Can the diagnosed influence be attributed to the assimilated radiosondes or do other observations also affect the tropopause structure?

The comparison of increments for 497 non-operationally launched radiosondes within an OSE confirms that the diagnosed influence (sharpness and altitude of the tropopause, wind acceleration in UT) can be mainly attributed to the assimilated radiosondes. However, the non--zero increments in the run without the NAWDEX radiosondes reveal that other observations also contributed to the sharpening and to the increase of wind at the radiosonde locations. The novel approach of a tropopause- -relative assessment in observations space combined with an OSE complements previous studies by providing a novel perspective on the local influence of DA on the tropopause that allows a positive influence to be assigned to the assimilation of radiosonde observations. Although the influence on the temperature and wind profiles is found to be small compared to the

27

645 background and analysis errors, DA is able to improve the sharp gradients of temperature and wind at the tropopause. The increased vertical gradients of temperature and wind are expected to improve the tropopause PV distribution (as indicated in Lavers et al., 2023). The sharpening process likely counteracts the decreasing forecast PV gradients. Future increases of horizontal and vertical model resolution in NWP and improved parameterizations of processes that modify the tropopause sharpness may positively impact the representation of the tropopause structure and thus the quality of NWP.

Data availability

650 The feedback files analysed in this study are requested from via the Meteorological Archival and Retrieval System (MARS) documented at https://confluence.ecmwf.int/display/UDOC/MARS+user+documentation, last access: 1st June 2023. We are grateful to the ECMWF for granting the data access.

Author contributions

655 KK performed the data analysis, produced the figures and wrote the manuscript. AS retrieved the observation feedback files. KK, AS, MW and GC conceptualized the study. MW and GC gave important guidance for the study, helped with the interpretation of the results and commented on the paper.

Competing interests

The authors declare that they have no conflicts of interest.

References

~~Bannister, R. N.: A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances, Q. J. Roy. Meteorol. Soc., 134, 1951–1970, https://doi.org/10.1002/qj.339, 2008.~~

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

Birner, T.: Fine-scale structure of the extratropical tropopause region, J. Geophys. Res., 111, D04104, https://doi.org/10.1029/2005JD006301, 2006.

Birner, T., Sankey, D., and Shepherd, T. G.: The tropopause inversion layer in models and analyses, Geophys. Res. Lett., 33, L14804, https://doi:10.1029/2006GL026549, 2006.

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. R. Meteorol. Soc., 140, 3878–3898, https://doi.org/10.1002/qj.4150, 2021.

Boljka, L. and Birner, T.: Potential impact of tropopause sharpness on the structure and strength of the general circulation. npj Clim. Atmos. Sci. 5, 98, https://doi.org/10.1038/s41612–022–00319–6, 2022.

Bonavita, M.: On some aspects of the impact of GPSRO observations in global numerical weather prediction, Q. J. Roy. Meteor. Soc., 140, 2546–2562, https://doi.org/10.1002/qj.2320, 2014.

Cucurull, L., and Anthes, R. A.: Impact of infrared, microwave and radio occultation satellite observations in operational numerical weather prediction, Mon. Weather Rev., 142, 4164–4186, https://doi.org/10.1175/MWR-D-14-00101.1, 2014.

Borne, M., Knippertz, P., Weissmann, M., Martin, A., Rennie, M. and Cress, A.: Impact of Aeolus wind lidar observations on the representation of the West African monsoon circulation in the ECMWF and DWD forecasting systems, Q. J. Roy. Meteor. Soc., 149, 933–98, https://doi.org/10.1002/qj.4442, 2023.

ECMWF, 2016: IFS Documentation – Cy43r1: Part III: Dynamics and Numerical Procedures, IFS Documentation, ECMWF, https://doi.org/10.21957/m1u2yxwrl~~https://www.ecmwf.int/en/elibrary/79987–ifs–documentation–cy43r1–part–ii–data–assimilation~~, last access: 1st December~~June~~ 2023, 2016.

29

705

Ferreira, A. P., Castanheira, J. M., and Gimeno, L.: Water vapour stratification and dynamical warming behind the sharpness of the Earth's midlatitude tropopause, Q. J. Roy. Meteor. Soc., 142, 957–970, https://doi.org/10.1002/qj.2697, 2016.

Gettelman, A., Hoor, P., Pan, L. L., Randel, W. J., Hegglin, M. I., and Birner, T.: The extratropical upper troposphere and
710 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.

715 Grise, K. M., Thompson, D. W. J., and Birner, T.: A global survey of static stability in the stratosphere and upper troposphere, J. Climate, 23, 2275–2292, https://doi.org/10.1175/2009JCLI3369.1, 2010.

Haualand, K. F. and Spengler, T.: Relative importance of tropopause structure and diabatic heating for baroclinic instability, Weather Clim. Dynam., 2, 695–712, https://doi.org/10.5194/wcd–2–695–2021, 2021.
720

Harvey, B., Methven, J., and Ambaum, M. H. P.: An Adiabatic Mechanism for the Reduction of Jet Meander Amplitude by Potential Vorticity Filamentation, J. Atmos. Sci., 75, 4091–4106, https://doi.org/10.1175/JAS-D-18-0136.1, 2018.

Hodyss, D. and Nichols, N.: The error of representation: Basic understanding, Tellus A, 67, 1–17,
725 https://doi.org/10.3402/tellusa.v67.24822, 2015.

Hoerling, M. P., Schaack, T. K., and Lenzen, A. J.: Global Objective Tropopause Analysis, Mon. Weather
Rev., 119, 1816–1831, https://doi.org/10.1175/1520-0493(1991)119<1816:GOTA>2.0.CO;2, 1991.

730 Hoffmann, L. and Spang, R.: An assessment of tropopause characteristics of the ERA5 and ERA-Interim meteorological reanalyses, Atmos. Chem. Phys., 22, 4019–4046, https://doi.org/10.5194/acp-22-4019-2022, 2022.

Houchi, K., Stoffelen, A., Marseille, G. J., and De Kloe, J.: Comparison of wind and wind shear climatologies derived from high-resolution radiosondes and the ECMWF model, J. Geophys. Res.-Atmos., 115, D22123,
735 https://doi.org/10.1029/2009JD013196, 2010.

Ingleby, B., Pauley, P., Kats, A., Ator, J., Keyser, D., Doerenbecher, A., Fucile, E., Hasegawa, J., Toyoda, E., Kleinert, T., Qu, W., St. James, J., Tennant, W., and Weedon, R.: Progress toward high-resolution, real-time radiosonde reports, B. Am. Meteorol. Soc., 97, 2149–2161, https://doi.org/10.1175/BAMS–D–15–00169.1, 2016.

740

Ingleby, B.: An assessment of different radiosonde types 2015/2016, ECMWF Technical Memorandum 69, https://doi.org/10.21957/0nje0wpsa, 2017.

Janjic, T., Bormann, N., Bocquet, M., Carton, J. A., Cohn, S. E., Dance, S. L., Losa, S. N., Nichols, N. K., Potthast, R., Waller, J. A., and Weston, P.: On the representation error in data assimilation, Q. J. R. Meteorol. Soc., 144, 1257–1278, https://doi.org/10.1002/qj.3130, 2018.

745

Kaluza, T., Kunkel, D., and Hoor, P.: On the occurrence of strong vertical wind shear in the tropopause region: a 10 year ERA5 northern hemispheric study, Weather Clim. Dynam., 2, 631–651, https://doi.org/10.5194/wcd-2-631-2021, 2021.

750

Krüger, K., Schäfler, A., Wirth, M., Weissmann, M., and Craig, G. C.: Vertical structure of the lower-stratospheric moist bias in the ERA5 reanalysis and its connection to mixing processes, Atmos. Chem. Phys., 22, 15559–15577, https://doi.org/10.5194/acp–22–15559–2022, 2022.

755   König, N., Braesicke, P., and von Clarmann, T.: Tropopause altitude determination from temperature profile measurements of reduced vertical resolution, Atmos. Meas. Tech., 12, 4113–4129, https://doi.org/10.5194/amt–12–4113–2019, 2019.

Lavers, D. A., Torn, R. D., Davis, C., Richardson, D. S., Ralph, F. M., and Pappenberger, F.: Forecast evaluation of the North Pacific jet stream using AR RECON Dropwindsondes, Q. J. Roy. Meteorol. Soc., Accepted Author Manuscript, 760   https://doi.org/10.1002/qj.4545, 2023.
Lavers, D. A., Torn, R. D., Davis, C., Richardson, D. S., Ralph, F. M., and Pappenberger, F.: Forecast evaluation of the North Pacific jet stream using AR Recon Dropwindsondes, Q. J. Roy. Meteorol. Soc., 149, 3044–3063, https://doi.org/10.1002/qj.4545, 2023.

765   Martin, A., Weissmann, M., and Cress, A.: Investigation of links between dynamical scenarios and particularly high impact of Aeolus on numerical weather prediction (NWP) forecasts, Weather Clim. Dynam., 4, 249–264, https://doi.org/10.5194/wcd-4-249-2023, 2023.

Martius, O., Schwierz, C., and Davies, H. C.: Tropopause-Level Waveguides, J. Atmos. Sci., 67, 866–879, 770   https://doi.org/10.1175/2009JAS2995.1, 2010.

31

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.

775   Pauley, P. and Ingleby, B.: Assimilation of in-situ observations, in: Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. IV), edited by: Park, S. K. and Xu, L., Springer, ISBN 978–3–030–77722–7, available at: https://link.springer.com/book/10.1007/978–3–030–77722–7, (last access: 1 June 2023), 2022.

Pilch Kedzierski, R., Matthes, K., and Bumke, K.: Synoptic-scale behavior of the extratropical tropopause inversion layer,
780   Geo-phys. Res. Lett., 42, 10018–10026, https://doi.org/10.1002/2015GL066409, 2015.

Pilch Kedzierski, R., Neef, L., and Matthes, K.: Tropopause sharpening by data assimilation, Geophys. Res. Lett., 43, 8298–8305, https://doi.org/10.1002/2016GL069936, 2016.

785   Polichtchouk, I., Bechtold, P., Bonavita, M., Forbes, R., Healy, S., Hogan, R., Laloyaux, P., Rennie, M., Stockdale, T., Wedi, N., Diamantakis, M., Flemming, J., English, S., Isaksen, L., Vána, F., Gisinger, S., and Byrne, N.: Stratospheric Modelling and Assimilation, ECMWF Technical Memorandum 877, https://doi.org/10.21957/25hegfoq, 2021.

Radnóti, G., Bauer, P., McNally, A., Cardinali, C., Healy, S., and de Rosnay, P.: ECMWF study on the impact of future
790   developments of the space-based observing system on numerical weather prediction. ECMWF Technical Memorandum Tech. Memo. 638, https://doi.org/10.21957/skfhvask, 2010.

Rabier, F., Järvinen, H., Klinker, E., Mahfouf, J.–F., and Simmons, A.: The ECMWF operational implementation of four–dimensional variational assimilation. I: Experimental results with simplified physics, Q. J. Roy. Meteorol. Soc., 126, 1148–
795   1170, https://doi.org/10.1002/qj.49712656415, 2000.

Saffin, L., Gray, S. L., Methven, J., and Williams, K. D.: Processes Maintaining Tropopause Sharpness in Numerical Models, J. Geophys. Res.–Atmos., 122, 9611–9627, https://doi.org/10.1002/2017JD026879, 2017.

800   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.,

805 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,
810 2889–2907, https://doi.org/10.1175/MWR–D–19–0229.1, 2020.

Schindler, M., Weissmann, M., Schäfler, A., and Radnóoti, G.: The impact of dropsonde and extra radiosonde observations during NAWDEX in autumn 2016, Mon. Weather Rev., 148, 809–824, https://doi.org/10.1175/MWR–D–19–0126.1 , 2020.

815 Schmidt, T., Wickert, J., and Haser, A.: Variability of the upper troposphere and lower stratosphere observed with GPS radio occultation bending angles and temperatures, Adv. Space Res., 46, 150–161, https://doi.org/10.1016/j.asr.2010.01.021 , 2010.

Schwierz, C., Dirren, S., and Davies, H. C.: Forced Waves on a Zonally Aligned Jet Stream, J. Atmos. Sci., 61, 73–87, https://doi.org/10.1175/1520–0469(2004)061<0073:FWOAZA>2.0.CO;2, 2004.
820
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.

Tinney E. N., Homeyer C. R., Elizalde L., Hurst D. F., Thompson A. M., Stauffer R. M., Vömel H. and Selkirk H. B.: A
825 modern approach to stability–-based definition of the tropopause. Mon. Weather Rev., 150, 3151–3174, https://doi.org/10.1175/MWR–D–22–0174.1, 2022.

Vaisala, 2017: Comparison of Vaisala Radiosondes RS41 and RS92, retrieved via: https://www.vaisala.com/sites/default/files/documents/RS-Comparison-White-Paper-
830 B211317EN.pdfhttps://www.vaisala.com/sites/default/files/documents/RS–Comparison–White–Paper–B211317EN.pdf, date of last access: 1st JulyDecember, 2023.

Weissmann, M., R. Busen, A. Dörnbrack, S. Rahm, and O. Reitebuch, 2005: Targeted observations with an air-borne wind lidar, J. Atmos. Oceanic Technol., 22, 1706-1719, https://doi.org/10.1175/JTECH1801.1.
835
Weissmann, M. and Cardinali, C.: Impact of airborne Doppler lidar observations on ECMWF forecasts, Q. J. Roy. Meteorol. Soc., 133, 107–116, https://doi.org/10.1002/qj.16, 2007.

Weissmann, M., Langland, R. H., Cardinali, C., Pauley, P. M., and Rahm, S.: Influence of airborne Doppler wind lidar profiles
840   near Typhoon Sinlaku on ECMWF and NOGAPS forecasts, Q. J. Roy. Meteorol. Soc., 138, 118–130,
https://doi.org/10.1002/qj.896, 2012.

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