https://doi.org/10.5194/acp-2020-254
Preprint. Discussion started: 14 April 2020

Atmospheric Chemistry and Physics Discussions

Using a global network of temperature lidars to identify temperature biases in the upper stratosphere in ECMWF reanalyses

Graeme Marlton[1], Andrew Charlton-Perez[1], Giles Harrison[1], Inna Polichtchouk[2], Alain Hauchecorne[3], Philippe Keckhut[3], and Robin Wing[3]

[1]Department of Meteorology, University of Reading, Reading, RG6 6LA
[2]European Centre for Medium Range Weather Forecasts, Shinfield Road, Reading, United Kingdom
[3]LATMOS/IPSL, UVSQ Université Paris-Saclay, Sorbonne Univerités, CNRS, Guyancourt, France

Correspondence: Graeme Marlton (graeme.marlton@reading.ac.uk)

Abstract. To advance our understanding of the stratosphere, high quality observational datasets of the upper atmosphere are needed. It is commonplace that reanalysis is used to conduct stratospheric studies. However the accuracy of the standard reanalysis at these heights is hard to infer due to a lack of in-situ measurements. Satellite measurements provide one source of temperature information. As some satellite information is already assimilated into reanalyses, the direct comparison of satellite temperatures to the reanalysis is not truly independent. Stratospheric lidars use Rayleigh scattering to measure density in the upper atmosphere, allowing temperature profiles to be derived for altitudes from 30km (where Mie scattering due to stratospheric aerosols becomes negligible) to 80-90km (where the signal-to-noise begins to drop rapidly). The Network for the Detection of Atmospheric Composition Change (NDACC) contains several lidars at different latitudes that have measured atmospheric temperatures since the 1970s, resulting in a long running upper-stratospheric temperature dataset. These temperature datasets are useful for validating reanalysis datasets in the stratosphere, as they are not assimilated into reanalyses. Here we take stratospheric temperature data from lidars in the northern hemisphere for winter months between 1990-2017 and compare them with the European Centre for ECMWF's ERA-interim and ERA-5 reanalyses. To give confidence in any bias found, temperature data from NASA's EOS Microwave Limb Sounder is also compared to ERA-interim and ERA-5 at points over the lidar sites. In ERA-interim a cold bias of -3 to -4 K between 10 hPa and 1 hPa is found when compared to both measurement systems. Comparisons with ERA-5 found a small bias of magnitude 1 K which varies between cold and warm bias with height between 10 hPa and 3 hPa, indicating a good thermal representation of the upper atmosphere to 3 hPa. At heights above this, comparisons with EOS MLS yield a slight warm bias and the temperature lidar yield a cold bias. A further comparison is undertaken to see the effects of the assimilation of the Advanced Microwave Sounding Unit-A satellite data and the Constellation Observing System for Meteorology, Ionosphere, and Climate GPS Radio Occulation (COSMIC GPSRO) data on stratospheric temperatures. By comparing periods before and after the introduction of each data source it is clear that COSMIC GPSRO improves the cold bias in the 3 hPa to 0.5 hPa altitude range

https://doi.org/10.5194/acp-2020-254
Preprint. Discussion started: 14 April 2020

Atmospheric Chemistry and Physics Discussions

1 Introduction

The middle atmosphere influences the weather and climate in the troposphere e.g Domeisen (2019); Domeisen et al. (2019). Sudden Stratospheric Warmings (SSWs) can cause changes in the tropospheric flow for many weeks e.g. Charlton and Polvani (2007) and the Quasi Biennial Oscillation (QBO), a 28 month switching in equatorial stratospheric winds, also affects the large scale processes in the troposphere e.g. Baldwin et al. (2001). Critical to our understanding of how these processes work are good observational datasets of the upper atmosphere. Often reanalysis datasets such as the European Centre for Medium range Weather Forecasts' (ECMWF) ReAnalaysis (ERA) datasets (Dee et al., 2011) or the US National Centre for Atmospheric Research (NCAR) reanalysis (Kalnay et al., 1996) amongst many others are used for stratospheric studies on a global scale as shown for example in Fujiwara et al. (2017); Seviour et al. (2012); Lee et al. (2019); Skerlak et al. (2014); Butler et al. (2015).

To create reanalysis datasets, a large amount of temperature data is assimilated from satellite and in-situ measurements. In the middle and upper stratosphere the number of temperature observations is somewhat limited. This makes diagnosing bias in a reanalysis dataset more difficult. Radiosondes, small balloon borne instrument packages, which provide in-situ temperature profiles up to heights of 30 to 40 km are launched from thousands of locations daily, giving a wealth of information that is assimilated in the lower and middle atmosphere. Simmons et al. (2020) undertook a study examining the performance of ERA-5, the ECMWF's most recent reanalysis dataset using radiosonde data. However there are height limitations, and the technique has the potential to be biased due to the assimilation of radiosonde observations within the reanalysis. Rocketsondes (Schmidlin, 1981) can reach heights of 100 km providing high resolution temperature information in the middle and upper atmosphere. Rocketsondes are not operationally assimilated due to the significant installations and costs to launch, which leads to sparse temporal and spatial sampling.

There are numerous satellite techniques to retrieve temperature profiles of the upper atmosphere. Stratospheric sounding units (SSU) (Miller et al., 1980) derive temperature from radiances in CO_2 emissions. Similarly the Sounding of the Atmosphere using Broadband Emission Radiometry (SABER) instrument uses limb emissions from CO_2 to provide temperature observations in the mesosphere and thermosphere (Russell et al., 1999). The Aqua satellite combines data from Atmospheric Infra-red Sounders (AIRS) with data from the Advanced Microwave Sounding Units (AMSU) to provide temperature profiles in the troposphere and stratosphere with 1 km vertical resolution (Susskind et al., 2006). The Microwave Limb Sounder (MLS) provides temperature data by observing the limb emission of several atmospheric gases and aerosols (Waters et al., 2006). Low Earth orbiters such as those in the Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC) can derive atmospheric properties such as temperature, pressure and water vapour using GPS Radio Occulation (GPSRO) up to 40 km (Kuo et al., 2001). As COSMIC is a constellation of satellites it retrieves thousands of randomly sampled temperature profiles daily across the globe. Many such techniques are assimilated into reanalysis making it hard to make an unbiased comparison.

Another source of middle atmosphere temperature measurements is from Rayleigh temperature lidars. These temperature lidars use Rayleigh scattering properties of the atmosphere above the stratospheric aerosol layer (>30 km) to infer density and, by assuming hydrostatic equilibrium, temperature (Hauchecorne and Chanin, 1980). There are several Rayleigh lidars across

https://doi.org/10.5194/acp-2020-254
Preprint. Discussion started: 14 April 2020


Atmospheric Chemistry and Physics Discussions

the globe based at participating sites in the Network for the Detection of Atmospheric Composition Change (NDACC). A small handful have been making temperature profile measurements between 30 km and 90 km for at least three decades. Furthermore the lidar temperature profiles are not assimilated into reanalysis making them independent for numerical dataset comparisons. Le Pichon et al. (2015) compared 6 months of Rayleigh temperature lidar data with ECMWF reanalysis data and found good agreement.

In this paper, we use temperature data from four lidars and compare it to temperature data from the ECMWF's ERA-interim and ERA5 reanalysis packages over three decades to identify temperature bias for stratospheric studies. To add confidence to the identified bias the same comparison will also be undertaken with temperature data from the National Aeronautics and Space Administrations (NASA) Earth Observing System Microwave Limb Sounder (EOS MLS). EOS MLS is one of the few satellite temperature datasets which is not assimilated into ECMWF reanalysis.

The reanalysis packages cover long time periods over which the quantity and types of data assimilated has increased. Poli et al. (2010) showed that inclusion of GPSRO data improved temperature bias in the upper stratosphere and lower troposphere. Hence, further analysis will be undertaken comparing ERA-interim and ERA5 before and after the inclusion of GPSRO data from the COSMIC constellation at the end of 2006. Simmons et al. (2020) also found that the inclusion of AMSU-A data in 2000 caused a increase in cold bias in ERA-5 we examine this here also.

2 Dataset description

2.1 Stratospheric Temperature Lidar

NDACC contains a network of approximately 30 stratospheric lidars distributed across the globe. Twelve of these lidars have the ability to observe temperature profiles in the upper atmosphere using Rayleigh backscattering methods. When a laser pulse is emitted into the atmosphere a small fraction of the photons are backscattered by both air molecules and aerosols. Above the stratospheric aerosol layer, at approximately 30 km, the dominant backscatter mechanism is from Rayleigh scattering of air molecules, which is proportional to the density of the atmosphere. From the retrieved density profile the hydrostatic balance and the ideal gas law are used to retrieve a temperature profile as described in depth in Hauchecorne and Chanin (1980). Here, a brief summary is given. A model atmosphere such as the Committee on Space Research's International Reference atmosphere (CIRA)-86 (Fleming et al., 1988) or the Naval Research Laboratory's Mass Spectrometer and Incoherent Scattering Radar model (NRLMSISE-00) (Picone et al., 2002) is first needed. Both are a function of time of year and latitude, and are used to infer the pressure of the uppermost edge of the lidar's highest observable range gate $P_{k+1/2}$. The highest observable range gate is defined as that where the signal to noise ratio is equal to one. As density ρ at this height is then known the pressure at the lower edge of the range gate $P_{k-1/2}$ can be inferred using the hydrostatic balance

$$P_{k-1/2} = P_{k+1/2} + \rho g(z) \Delta z, \tag{1}$$

where Δz is the range gate depth and $g(z)$ is the acceleration due to gravity which is typically 9.81 m s^{-2} at the surface. However, g decreases with height, at an altitude of 50 km the acceleration due to gravity decreases by approximately 1%

https://doi.org/10.5194/acp-2020-254
Preprint. Discussion started: 14 April 2020


Atmospheric Chemistry and Physics Discussions

compared to that at the Earth's surface. Further to this $g(z)$ also has latitudinal variability due to the rotating oblique spheroidal nature of the Earth. P across the centre of the range gate can be calculated by averaging $P_{k+1/2}$ and $P_{k-1/2}$, which can be combined with ρ to calculate the temperature T across the range gate using the perfect gas law

$$T = \frac{P}{\rho R} \tag{2}$$

where R is the gas constant with a value of 287 J kg^{-1}. Thus, through numerical integration downwards the pressure and temperature of the lidar's density profile can be inferred. However, as the uppermost pressure value is estimated from a model which does not account for upper atmosphere variability the initial pressure level guess may incur a large error. Experimental work by Leblanc et al. (1998) and Jalali et al. (2018) found that the uncertainties could be of the order 10 K in the highest 10 km of the derived temperature profile if the real atmosphere differed largely from the climate atmosphere. For this reason the top 10 km of the profiles are often excluded from published datasets (Wing et al., 2018b). One drawback of Rayleigh lidar is that they can only by used effectively at night so the returned lidar signal is not lost in the interference from solar radiation. Optical filtering methods can be used to remove the interference from solar radiation. However, they are expensive and the height from which temperature retrievals can be made is much reduced. In order to make retrievals the sky must also be cloud free so the beam does not become attenuated.

Calibrating Rayleigh temperature lidars can be difficult due to a lack of other temperature observations at these altitudes. One method undertaken to validate a lidar's measurements is to use a mobile benchmark lidar, which has been extensively characterised (Keckhut et al., 2004), and can be transported from site to site to blind test each lidar (Steinbrecht et al., 2009). The current benchmark lidar is operated by the NASA Goddard Space Flight Centre (GSFC) (McGee et al., 1995) and has been used to compare the temperature lidars at Mauna Loa, Observatory Haute Provence, Table Mountain Observatory (Keckhut et al., 2004) and Hohenpeissenburg (Steinbrecht et al., 2009). The NASA GSFC lidar has been compared with a combination of different independent technologies in the Stratospheric Ozone Inter-comparison (STOIC) campaign (Ferrare et al., 1995) in which the lidar was compared to radiosondes, rocketsondes and satellite data. Comparison with radiosondes between 35 km and 40 km found the lidar observed temperatures 1-2 K warmer. Given the upper operating limit of these radiosondes the 1-2 K bias falls within the estimated $\pm$4 K error of the radiosonde (Ferrare et al., 1995). In comparison with the rocketsondes the NASA GSFC lidar recorded colder temperatures in the same height range with a typical difference of -2 to -3 K. The precision of the rocketsondes from repeatability experiments to two standard deviations between 30-40 km was found to be $\pm$ 1 K and decreased to $\pm$1.8 K at 55 km, and above 60 km the precision decreases to $\pm$ 3.3 K (Schmidlin, 1981). Ferrare et al. (1995) comments that given the instrumental accuracies and the co-location and temporal differences between lidar and rocket temperature soundings, there is a reasonable level of agreement between the two. A comparison was also made with the Stratospheric Aerosol and Gas Experiment (SAGE II) satellite which found that for the STOIC campaign period there was large variability likely due to distances between the satellite scans and the campaign site. Another experiment which compared lidars at OHP to the SAGE II temperature measurements (Wang et al., 1992) found the satellite and lidar showed good agreement with the satellite (within 2 K between 30 and 55 km). From these experiments it is evident that the NASA GSFC lidar is

https://doi.org/10.5194/acp-2020-254
Preprint. Discussion started: 14 April 2020


Atmospheric Chemistry and Physics Discussions

accurate to ± 2K over the atmosphere range of 30-60 km. This means that the lidar chosen for this study should also carry an accuracy similar to ± 2 K, due to the comparisons made with the NASA GFC lidar.

Other sources of error have been attributed to range gate height errors and altitude shifts (Steinbrecht et al., 2009). There are also potential sources of uncertainty due to the presence of gravity waves in the temperature profiles. Jenkins et al. (1987) showed that large amplitude gravity waves with temperature perturbations in the region of 20 K could produce an error of ± 1 K if the observation time was less than the period of the gravity wave. In many cases the lidars observe for several hours and the temperature perturbations of the wave are removed through averaging leading to negligible error.

In summary, methods of obtaining temperature profiles using lidars are robust. In comparisons with other in-situ and remote sensing measurement technologies biases are in the range of ± 2K over this height range.

For the comparisons being undertaken here lidar temperature datasets are needed which span at least 10 years and have frequent temperature profiles during that period. Of the 12 lidar sites within NDACC, only four lidar datasets fitted this criteria and are shown in table 1. It should be further noted that both the Mauna-Loa and Table Mountain observatory lidars have Raman channels. The stratospheric aerosol layer does not affect the Raman scattering of Nitrogen to the same degree as Rayleigh scattering and thus allows an atmospheric temperature profile to be inferred in a near similar method to that using Rayleigh scattering (Gross et al., 1997). This allows the Table Mountain Observatory and Mauna Loa lidars to observe temperatures to approximately 8 km. For these two lidars the reanalysis packages described in section 2.3 will be compared for greater altitude ranges.

2.2 Microwave Limb Sounder

The NASA EOS MLS was launched on the 14th July 2004 and became operational on the 14th August 2004. It works by observing millimetre and sub-millimetre thermal emission along the limb of the atmosphere. It is in a low polar orbit allowing it to orbit the earth 15 times a day, crossing the equator near to local noon and midnight. It is designed to measure several atmospheric gases and aerosols in the upper troposphere, stratosphere and mesosphere (Waters et al., 2006). It uses the emission from oxygen to provide temperature and pressure measurements; the precision of the measurement is given in Waters et al. (2006) to be 0.5-1 K between 300 and 0.001 hPa, where it has a vertical resolution between 4-8 km. Initial comparisons by Froidevaux et al. (2006) with other satellite retrievals of temperature found that EOS MLS had a 1-2 K warm bias. A more thorough comparison made by Schwartz et al. (2008) compared EOS MLS temperature retrievals to those from radiosondes, several satellites, and from GPSRO. It was found that from 0.0001 hPa to 0.3 hPa that temperature bias could range from -9 to 0K with temperature precision ranging from ±1 to ±2.5 K. A further study by Wing et al. (2018a) found that the bias in wintertime MLS was ±10 K and ±4 K in the summertime. At 1 hPa warm biases from 0 to 5K were found. From 3.16 hPa down to 316 hPa precision was found to be less than ±1 K and biases were between -2 and 1.5 K. Thus, at pressure heights of 3 hPa and below the EOS MLS satellite has a similar bias to that of the temperature lidar at the same observing height.

https://doi.org/10.5194/acp-2020-254
Preprint. Discussion started: 14 April 2020


Atmospheric Chemistry and Physics Discussions

2.3 European Centre for Medium Range Forecasting data

A reanalysis dataset is generated by combining many different historical measurements using data assimilation to create an accurate numerical representation of the Earth's atmosphere at a given time. ERA interim spans from 1979 to present and at the time of writing runs 2-3 months behind real time. It is described in Dee et al. (2011) and has 60 model levels with an approximate 79 km horizontal grid resolution and 6 hour analysis windows and is based on the Integrated Forecasting System (IFS) cycle 31R2. Due to the temporal range of the ERA-interim reanalysis we are able to compare the entire temperature lidar datasets shown in table 1.

The ERA-5 is the 5th generation reanalysis created by the ECMWF and replaces ERA-interim. The new ERA-5 reanalysis, based on IFS cycle 41r2 (ECMWF, 2016) has 137 model levels from the surface to 0.01 hPa and a global horizontal resolution of 31 km, compared to ERA Interim's 60 model levels and 79 km horizontal resolution. Other improved features are the inclusion of more measurements and improved bias correction techniques to assimilate them and better climate forcings (Hersbach and Dee, 2016). Dee et al. (2011) state that during December 2006 GPSRO data from the COSMIC constellation was included in reanalysis datasets and that the variability in these datasets was much improved after this time (Poli et al., 2010). For the second part of the analysis, comparisons with both ERA-interim and ERA-5 before and after the inclusion of COSMIC GPSRO data will be undertaken. In a previous study by Poli et al. (2010) the comparison was made between 200 hPa and 100 hPa. Here the comparison will be made at 100 hPa and above.

3 ERA-interim comparisons

In this section MLS and lidar profiles will be compared with temperature profiles from ERA-interim. For all of our comparisons we will only focus on the Northern Hemisphere winter months between October and March. Stratospheric variability increases during these months and we examine whether this variability is present within the reanalyses. The lidar temperature profiles were interpolated onto ERA-interim's model levels using Z, the geopotential height, from the reanalysis package for time steps that were closest to the midpoint of the lidar's profile acquisition period. The height resolution for ERA-interim at these heights is approximately 1.5 km at 10 hPa and 3 km at 1 hPa. For this comparison we use lidar and reanalysis profiles for the full temporal lidar temperature profile range shown in table 1. Figure 1 panels (a-d) shows matched mean temperature profiles from both lidar (red) and ERA-interim (blue) for the lidar sites at Hohenpeissenburg, Mauna-Loa, OHP and Table Mountain Observatory respectively. Panels e-h show the mean of the matched differences for the corresponding lidars a-d. ERA-interim at the points studied here has a cold bias in the region of -3 to -4 K. Red dots show where the mean difference is significantly different from zero at the 95% significance level. For Hohenpeissenburg, OHP and Table Mountain Observatory, the peak cold bias is centred between 10hPa and 1hPa. For Mauna Loa the cold bias extends down to the 100 hPa pressure surface. Reasons for this could be due to Hohenpeissenburg and Table Mountain Observatory being at higher latitudes than the tropically positioned Mauna-Loa station where the structure of the middle atmosphere differs slightly. In all cases above 1 hPa ERA-interim is much warmer as it approaches 0.1 hPa.

https://doi.org/10.5194/acp-2020-254
Preprint. Discussion started: 14 April 2020


Atmospheric Chemistry and Physics Discussions

A similar analysis was performed using the EOS MLS data. The EOS MLS temperature data was first sorted so that only night time passes within 2.5^o of each temperature lidar were retained. The remaining EOS MLS temperature profiles were interpolated on to ERA-interim's model levels using Z from the reanalysis package for time steps that were closest and less than 3 hours apart. Due to EOS MLS only being launched in 2004 the results of the ERA-interim and MLS comparison shown in figure 2 are for the years 2004 to 2017. Figure 2 shows matched mean temperature profiles from both MLS (red) and ERA-interim (blue) for lidar sites at Hohenpeissenburg, Mauna-Loa, OHP and Table Mountain Observatory respectively. Panels (e-g) show the mean of the matched differences for MLS over the 4 lidar sites. The cold bias seen in figure 1 between 1 hPa and 10 hPa is still present. However the size of the bias is smaller with a cold bias of -1 to -2 K. For comparisons over Hohenpeissenburg and OHP (panels e and g) the cold bias detected by the MLS is reduced when compared to that of the lidar. Between 10 hPa and 100 hPa MLS observes a warm bias of 1 K in ERA-interim which cannot be verified by the temperature lidar at these two locations. For comparisons with Table mountain observatory and Mauna-Loa a cold bias is also observed from 1 hPa down to 100 hPa similar to the bias found by the lidar at these two sites. The only difference is that the magnitude of the bias is reduced by 1-2 K. This may be due to the warm bias of EOS MLS at these heights as shown in Schwartz et al. (2008). This demonstrates evidence of a systematic bias between the lidar and MLS measurement technologies. Simmons et al. (2020) compared the global mean temperature from ERA-interim with global mean radiosonde data between 15 hPa and 7 hPa and 7 hPa and above. Here our results agree with a cold bias seen in Simmons et al. (2020)'s at 7 hPa and above results. For the 15 hPa to 7 hPa range the bias differ, Simmons et al. (2020) found that there was a slight warm bias of 0.5 K. Figure 1 panels (e) and (h) and figure 2 panels (e,g and h) show evidence towards this.

In brief conclusion two independent observation methods have shown that a cold bias exists between 1 and 10 hPa in ERA-interim over Hohenpeissenburg and OHP lidar and between 1 and 100 hPa for the Table Mountain observatory and Manua-Loa sites. The reason for the differnece in the magnitude of the bias is likely down to the differences in measurement technology. As discussed in section 2, the MLS has been found to exhibit a warm bias between 1 and 3 hPa (Schwartz et al., 2008), which may be the reason for a smaller cold bias than detected with the lidar.

4 ERA-5 Comparisons

To compare the temperature biases in ERA-5 with those in ERA-Interim we repeat the comparison in a similar manner. The temperature lidar data was interpolated onto ERA-5 model levels using Z for the nearest analysis window to the mid point of the lidar acquisition window. Figure 3 panels (a-d) show the mean temperature profiles of the lidar in red and ERA-5 in blue up to a height of 0.5 hPa for the winter months, October-March for the period 1990-2017. At first inspection ERA5 profiles track the lidar profiles more than those of ERA-interim (figure 1). Figure 3 panels (e-h) show the mean of matched differences with height; red dots show when the mean difference is different from zero at the 95% significance level inferred by a single sample t-test. The temperature biases are significantly smaller than in ERA-interim. For the Mauna-Loa and Table mountain sites the temperature bias is very close to zero up to the 3 hPa pressure surface. Above this height a cold bias, which increased to -5 K is observed at all four sites. OHP and Hohenpeissenburg show smaller temperature biases which vary by 1 K about zero. When

https://doi.org/10.5194/acp-2020-254
Preprint. Discussion started: 14 April 2020


Atmospheric Chemistry and Physics Discussions

we consider that the measurement accuracy of the lidar is ± 2 K, ERA5 gives a good thermal representation of the atmosphere up to 3 hPa.

Figure 4 shows the temperature comparisons against MLS for the period 2004 to 2017; the data was interpolated on to the ERA5 model data using the same method as discussed in section 3. MLS and ERA-5 show good agreement at all sites at all heights. Figure 4 panels (e-h) show that at the Hohenpeissenburg and OHP site there is a slight warm bias, which peaks at 1 K at 3 hPa. The temperature bias at Table Mountain Observatory (g) varies about zero but peaks at 3 hPa similar to the OHP and Hohenpeissenburg sites. The temperature bias at the Mauna Loa observatory in (f) varies about zero by approximately 0.5 K over the height range studied. The uncertainties in the bias are larger for MLS at higher altitudes which is shown by the increased breadth in the shading in figure 4 panels (e-h). This is likely due to the accuracies of MLS previously discussed. In addition, MLS was found to have a warm bias (Schwartz et al., 2008) between 0.3 hPa and 3 hPa which is seen here. Further to this MLS shows an oscillation with heights that is not seen in the lidar data. Simmons et al. (2020) compared the global mean temperature from ERA-5 with global mean radiosonde data from 15 hPa upwards and found that performance was similar to ERA-interim pre 2000. A cold bias of -2 K was found above 7 hPa and a slight warm bias of less than 0.5 K between 7-15 hPa for ERA-5. Between 2000 and 2007 there was an increase in the cold bias making the bias -3 K and -0.5 K for the layers 7 hPa and above and 7-15 hPa respectively. It was believed that this may be due to the introduction of the AMSU-A satellite data stream. Post 2007, after the introduction of GPSRO data, Simmons et al. (2020) show temperature bias above 7 hPa to be -1.5 K and less than -0.5 K between 7-15 hPa. As our comparisons using the lidar span 1990 to 2017 we do see some aspects of Simmons et al. (2020)'s results in that a cold bias is observed above 7 hPa and the temperature bias is small in the 7-15 hPa range. However, due to our methodology of averaging over 30 years for the lidar data it is hard to see if the increase in bias during 2000-2007 is present. Due to the time period of the MLS comparison such a comparison is not possible

MLS and ERA-5 temperatures agree well over the height ranges studied. The temperature lidar agrees up to 3 hPa with a cold bias observed above this. One reason MLS may not detect this cold bias is due to the fact it exhibits a warm bias, counteracting the cold bias at heights above 3 hPa (Schwartz et al., 2008). Aside from the two differing measurement techniques, the main difference in this study is that the lidar dataset is almost double the length of the EOS MLS dataset- temperature lidars have been used since the 1990s and EOS MLS only became operational in 2004. During this time period other measurement technologies have been added to those already assimilated by reanalysis datasets. It could be that the cold bias observed by the lidar above 3 hPa in ERA-5 is from an earlier time period in the dataset where observations were more sparse. We check this in the next section by examining temperature bias inferred by the lidar over three different time periods.

5 ERA performance due to Assimilation of Satellite data

In December 2006 the COSMIC constellation was launched and later that month GPSRO data was made available to assimilate into the ECMWF IFS models (Dee et al., 2011; Poli et al., 2010). Initial comparisons by Poli et al. (2010) showed that the inclusion of COSMIC GPSRO data improved model performance in the upper troposphere and lower stratosphere. However this comparison was only undertaken between 200 hPa and 100 hPa. Its effect at 100 hPa was described in Cardinali and Healy

https://doi.org/10.5194/acp-2020-254
Preprint. Discussion started: 14 April 2020


Atmospheric Chemistry and Physics Discussions

(2014) who found that the introduction of COSMIC GPSRO showed signs of improvement at higher levels. Also Simmons et al. (2020) showed a change in ERA-5 temperature bias at the inclusion of the AMSU-A satellite data. Here, an analysis is undertaken which will compare ERA-interim and ERA5 data to the lidar data both pre 1st January 2000, 1st January 2000 to 1st January 2007 and post 1st January 2007 to see how the introduction of AMSU-A and COSMIC GPSRO data effects upper stratosphere temperatures in ERA 5 and ERA-interim.

Figure 5 shows temperature difference between ERA-interim and the temperature lidar in this study for pre 2000 (green), 2000 to 2007 (purple) and post 2007 (cyan) data. The cold temperature bias at Hohenheissenburg (a) and OHP (c) in ERA-interim does not seem to affected significantly by either of AMSU-A or COSMIC GPSRO data. At Mauna-Loa between 5-10 hPa there is an increase in the cold bias from -1 to -2 K post 2000 which maybe due to the introduction of AMSU-A satellite. TMO in panel (d) of figure 6 is the only site which shows significant improvement due to the inclusion of AMSU-A and COSMIC GPSRO with a pre 2000 cold bias of -6 K and a post 2007 bias of -3 K.

Figure 6 shows the temperature difference between ERA-5 and the temperature lidar data in this study for pre 2000 (green), 2000 to 2007 (purple) and post 2007 (cyan) data. For Hohenpeissenburg, Mauna-loa and OHP the introduction of AMSU-A data appears to have little effect on the temperature bias at these heights. The only significant improvement is a reduction in warm bias at 5 hPa at Mauna-Loa. Table mountain in panel (d) shows significant differences- the cold bias above 3 hPa is reduced to under -5 K and the cold bias between 5 hPa and 10 hPa has switched to a warm bias of 1 K, an overall improvement. The temperature bias at TMO is further improved post 2007 with temperature biases of $\pm$ 1.5 K between 1 and 100 hPa. For the Hohenpeissenburg site, although the cold bias has been reduced above 3 hPa for the post 2007 periods, the cold bias in the 3-10 hPa range has increased. There are no significant changes between the 2000 and 2007 and post 2007 values for Mauna-Loa and for OHP the cold bias increases to -2 K post 2007.

In summary, the cold temperature bias in ERA5 above 3 hPa is reduced from 5 K to 1-3 K at 3 out of 4 of the sites post 2007 due to the inclusion of GPSRO and AMSU-A data. This demonstrates how a lengthy time set such as the temperature lidar can be used to assess the effect of the addition of different observation streams in reanalysis data sets in the upper atmosphere. When comparing these findings to that of Simmons et al. (2020) who showed an intensifying of the cold bias between 2000 and 2007 we see that in these comparisons this is not the case. The reasons for this may potentially be that the metric used in Simmons et al. (2020) was the global mean radiosonde temperature with a bias correction applied. However, the radiosonde data had been assimilated into ERA-5 already and this combined with the bias correction may lead to the cold bias shown. Alternatively and more likely is that as a global average was used, this may not represent temperature bias at individual sites, as is shown here. It is well known that temperature bias across the globe at the heights considered here is not homogeneous.

6 Conclusions

In this work we have utilised both temperature lidar from the NDACC network and EOS MLS data, that are not assimilated into models, to identify temperature bias in the upper atmosphere of ECMWF reanalysis datasets. For comparisons with ERA-interim and the lidar, a cold bias of -3 to -4 K between 10 hPa and 1 hPa and a large warm bias above 1 hPa was found.

https://doi.org/10.5194/acp-2020-254
Preprint. Discussion started: 14 April 2020

Atmospheric Chemistry and Physics Discussions

Comparisons with MLS also yielded cold biases between 10 hPa and 1 hPa but size of the cold bias was smaller at 2-3 K. Given the warm bias from MLS and that both measurement technologies have an approximate accuracy of $\pm$ 2 K a fair assessment of the cold bias in ERA-interim is between -4 to -2 K. Temperature bias between the lidar and ERA-5 were within $\pm$1 K to a height of 3 hPa and MLS showed similar agreement over the whole study height. Comparisons between ERA5 and MLS do hint at a slight warm bias at 3 hPa, although this is likely due to EOS MLS's bias at these height ranges. By splitting the lidar dataset into to three time periods, 1990-2000,2000-2007 and 2007 to 2017, it was found that the cold bias detected by the lidar above 3 hPa was reduced from 5 K to 2-3 K with the introduction of GPSRO in 2007.

From the comparisons here it can be stated that ERA-5 is much improved compared to ERA-interim and has a good thermodynamic representation of the upper stratosphere, when we consider the uncertainties in the lidar and MLS at these heights are $\pm$ 2 K, making it an excellent choice for further stratospheric studies or using it as reference with which to compare other models. However, a cold bias detected in ERA-5 by the lidar before the inclusion of GPSRO data should be accounted for.

Using both temperature lidar and a non-assimilated satellite such as EOS MLS, both have their advantages. EOS MLS has global coverage albeit coarser vertical resolution, and observations over a shorter time period. The temperature lidars on the other hand, whilst limited to a few locations globally, have high vertical resolution measurements that have been made for nearly 30 years, making them an important and useful tool for inferring temperature bias in reanalysis datasets. A future test could see the lidar networks used to explore modifications to reanalysis datasets such as testing the experimental ERA5.1 shown in Simmons et al. (2020)

Data availability. The temperature lidar data is available for public download through http://ndacc-lidar.org/index.php?id=70/Data.htm. The Microwave Limb Sounder data was available for public download at https://mls.jpl.nasa.gov/. ECMWF ERA-interim and ERA5 data are available from the ECMWF MARS archive.

Author contributions. GJM extracted the datasets, performed the analysis and prepared the manuscript ,RW processed the OHP data, ACP, RGH, IP, AH and PK provided inputs into the analysis and preparation of the manuscript.

Competing interests. There are no competing interests

Acknowledgements. This work was performed during the course of the ARISE2 collaborative infrastructure design study project funded by the European Commission H2020 program (grant number 653980, www.arise-project.eu).The authors also wish to acknowledge staff at the ECMWF for their discussions as this work advanced

https://doi.org/10.5194/acp-2020-254
Preprint. Discussion started: 14 April 2020

Atmospheric Chemistry and Physics Discussions

References

Baldwin, M., Gray, L., Dunkerton, T., Hamilton, K., Haynes, P., Randel, W., Holton, J., Alexander, M., Hirota, I., Horinouchi, T., et al.: The quasi-biennial oscillation, Reviews of Geophysics, 39, 179–229, 2001.

Butler, A. H., Seidel, D. J., Hardiman, S. C., Butchart, N., Birner, T., and Match, A.: Defining sudden stratospheric warmings, Bulletin of the American Meteorological Society, 96, 1913–1928, 2015.

Cardinali, C. and Healy, S.: Impact of GPS radio occultation measurements in the ECMWF system using adjoint-based diagnostics, Quarterly Journal of the Royal Meteorological Society, 140, 2315–2320, 2014.

Charlton, A. J. and Polvani, L. M.: A new look at stratospheric sudden warmings. Part I: Climatology and modeling benchmarks, Journal of Climate, 20, 449–469, 2007.

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P., et al.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, 2011.

Domeisen, D. I.: Estimating the frequency of sudden stratospheric warming events from surface observations of the North Atlantic Oscillation, Journal of Geophysical Research: Atmospheres, 124, 3180–3194, 2019.

Domeisen, D. I., Garfinkel, C. I., and Butler, A. H.: The teleconnection of El Niño Southern Oscillation to the stratosphere, Reviews of Geophysics, 57, 5–47, 2019.

ECMWF: Part III: Dynamics and Numerical Procedures, 2016.

Ferrare, R., McGee, T., Whiteman, D., Burris, J., Owens, M., Butler, J., Barnes, R., Schmidlin, F., Komhyr, W., Wang, P., et al.: Lidar measurements of stratospheric temperature during STOIC, Journal of Geophysical Research: Atmospheres, 100, 9303–9312, 1995.

Fleming, E. L., Chandra, S., Schoeberl, M. R., and Barnett, J. J.: Monthly mean global climatology of temperature, wind, geopotential height, and pressure for 0-120 km, 1988.

Froidevaux, L., Livesey, N. J., Read, W. G., Jiang, Y. B., Jimenez, C., Filipiak, M. J., Schwartz, M. J., Santee, M. L., Pumphrey, H. C., Jiang, J. H., et al.: Early validation analyses of atmospheric profiles from EOS MLS on the Aura satellite, IEEE Transactions on Geoscience and Remote Sensing, 44, 1106–1121, 2006.

Fujiwara, M., Wright, J. S., Manney, G. L., Gray, L. J., Anstey, J., Birner, T., Davis, S., Gerber, E. P., Harvey, V. L., Hegglin, M. I., et al.: Introduction to the SPARC Reanalysis Intercomparison Project (S-RIP) and overview of the reanalysis systems, Atmospheric Chemistry and Physics, 17, 1417–1452, 2017.

Gross, M. R., McGee, T. J., Ferrare, R. A., Singh, U. N., and Kimvilakani, P.: Temperature measurements made with a combined Rayleigh–Mie and Raman lidar, Applied Optics, 36, 5987–5995, 1997.

Hauchecorne, A. and Chanin, M.-L.: Density and temperature profiles obtained by lidar between 35 and 70 km, Geophysical Research Letters, 7, 565–568, 1980.

Hersbach, H. and Dee, D.: ERA5 reanalysis is in production, ECMWF Newsletter, 147, 2016.

Jalali, A., Sica, R. J., and Haefele, A.: Improvements to a long-term Rayleigh-scatter lidar temperature climatology by using an optimal estimation method., Atmospheric Measurement Techniques, 11, 2018.

Jenkins, D., Wareing, D., Thomas, L., and Vaughan, G.: Upper stratospheric and mesospheric temperatures derived from lidar observations at Aberystwyth, Journal of Atmospheric and Terrestrial Physics, 49, 287–298, 1987.

https://doi.org/10.5194/acp-2020-254
Preprint. Discussion started: 14 April 2020

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., et al.: The NCEP/NCAR 40-year reanalysis project, Bulletin of the American Meteorological Society, 77, 437–471, 1996.

Keckhut, P., McDermid, S., Swart, D., McGee, T., Godin-Beekmann, S., Adriani, A., Barnes, J., Baray, J.-L., Bencherif, H., Claude, H., et al.: Review of ozone and temperature lidar validations performed within the framework of the Network for the Detection of Stratospheric Change, Journal of Environmental Monitoring, 6, 721–733, 2004.

Kuo, Y.-H., Sokolovskiy, S., Anthes, R., and Vandenberghe, F.: Assimilation of GPS radio occultation data for numerical weather prediction, 2001.

Le Pichon, A., Assink, J., Heinrich, P., Blanc, E., Charlton-Perez, A., Lee, C. F., Keckhut, P., Hauchecorne, A., Rüfenacht, R., Kämpfer, N., et al.: Comparison of co-located independent ground-based middle atmospheric wind and temperature measurements with numerical weather prediction models, Journal of Geophysical Research: Atmospheres, 120, 8318–8331, 2015.

Leblanc, T., McDermid, I. S., Hauchecorne, A., and Keckhut, P.: Evaluation of optimization of lidar temperature analysis algorithms using simulated data, Journal of Geophysical Research: Atmospheres, 103, 6177–6187, 1998.

Lee, C., Smets, P., Charlton-Perez, A., Evers, L., Harrison, G., and Marlton, G.: The potential impact of upper stratospheric measurements on sub-seasonal forecasts in the extra-tropics, in: Infrasound Monitoring for Atmospheric Studies, pp. 889–907, Springer, 2019.

McGee, T. J., Ferrare, R. A., Whiteman, D. N., Butler, J. J., Burris, J. F., and Owens, M. A.: Lidar measurements of stratospheric ozone during the STOIC campaign, Journal of Geophysical Research: Atmospheres, 100, 9255–9262, 1995.

Miller, D., Brownscombe, J., Carruthers, G., Pick, D., and Stewart, K.: Operational temperature sounding of the stratosphere, Phil. Trans. R. Soc. Lond. A, 296, 65–71, 1980.

Picone, J. M., Hedin, A. E., Drob, D. P., and Aikin, A. C.: NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues, Journal of Geophysical Research: Space Physics, 107, SIA 15–1–SIA 15–16, https://doi.org/10.1029/2002JA009430, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2002JA009430, 2002.

Poli, P., Healy, S., and Dee, D.: Assimilation of Global Positioning System radio occultation data in the ECMWF ERA–Interim reanalysis, Quarterly Journal of the Royal Meteorological Society, 136, 1972–1990, 2010.

Russell, J. M., Mlynczak, M. G., Gordley, L. L., Tansock, J. J., and Esplin, R. W.: Overview of the SABER experiment and preliminary calibration results, in: Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research III, vol. 3756, pp. 277–289, International Society for Optics and Photonics, 1999.

Schmidlin, F. J.: Repeatability and measurement uncertainty of the United States meteorological rocketsonde, Journal of Geophysical Research: Oceans, 86, 9599–9603, 1981.

Schwartz, M., Lambert, A., Manney, G., Read, W., Livesey, N., Froidevaux, L., Ao, C., Bernath, P., Boone, C., Cofield, R., et al.: Validation of the Aura Microwave Limb Sounder temperature and geopotential height measurements, Journal of Geophysical Research: Atmospheres, 113, 2008.

Seviour, W. J., Butchart, N., and Hardiman, S. C.: The Brewer–Dobson circulation inferred from ERA-Interim, Quarterly Journal of the Royal Meteorological Society, 138, 878–888, 2012.

Simmons, A., Soci, C., Nicolas, J., Bell, B., Berrisford, P., Dragani, R., Flemming, J., Haimberger, L., Healy, S., Hersbach, H., Horányi, A., Inness, A., Munoz-Sabater, J., Radu, R., and Schepers, D.: Global stratospheric temperature bias and other stratospheric aspects of ERA5 and ERA5.1, 859, https://doi.org/10.21957/rcxqfmg0, https://www.ecmwf.int/node/19362, 2020.

Skerlak, B., Sprenger, M., and Wernli, H.: A global climatology of stratosphere-troposphere exchange using the ERA-Interim data set from 1979 to 2011, Atmospheric Chemistry and Physics, 14, 913, 2014.

https://doi.org/10.5194/acp-2020-254
Preprint. Discussion started: 14 April 2020

Atmospheric Chemistry and Physics Discussions

Steinbrecht, W., McGee, T., Twigg, L., Claude, H., Schönenborn, F., Sumnicht, G., and Silbert, D.: Intercomparison of stratospheric ozone and temperature profiles during the October 2005 Hohenpeißenberg Ozone Profiling Experiment (HOPE), Atmospheric Measurement Techniques, 2, 125–145, 2009.

Susskind, J., Barnet, C., Blaisdell, J., Iredell, L., Keita, F., Kouvaris, L., Molnar, G., and Chahine, M.: Accuracy of geophysical parameters derived from Atmospheric Infrared Sounder/Advanced Microwave Sounding Unit as a function of fractional cloud cover, Journal of Geophysical Research: Atmospheres, 111, 2006.

Wang, P.-H., McCormick, M., Chu, W., Lenoble, J., Nagatani, R., Chanin, M., Barnes, R., Schmidlin, F., and Rowland, M.: SAGE II stratospheric density and temperature retrieval experiment, Journal of Geophysical Research: Atmospheres, 97, 843–863, 1992.

Waters, J. W., Froidevaux, L., Harwood, R. S., Jarnot, R. F., Pickett, H. M., Read, W. G., Siegel, P. H., Cofield, R. E., Filipiak, M. J., Flower, D. A., et al.: The earth observing system microwave limb sounder (EOS MLS) on the Aura satellite, IEEE Transactions on Geoscience and Remote Sensing, 44, 1075–1092, 2006.

Wing, R., Hauchecorne, A., Keckhut, P., Godin-Beeckmann, S., Khaykin, S., and Mccullough, E.: Lidar temperature series in the middle atmosphere as a reference data set–Part 2: Assessment of temperature observations from MLS/Aura and SABER/TIMED satellites, Atmospheric Measurement Techniques, 11, 6703–6717, 2018a.

Wing, R., Hauchecorne, A., Keckhut, P., Godin-Beekmann, S., Khaykin, S., McCullough, E. M., Mariscal, J.-F., and d'Almeida, É.: Lidar temperature series in the middle atmosphere as a reference data set–Part 1: Improved retrievals and a 20-year cross-validation of two co-located French lidars, Atmospheric Measurement Techniques, 11, 5531–5547, 2018b.

https://doi.org/10.5194/acp-2020-254
Preprint. Discussion started: 14 April 2020

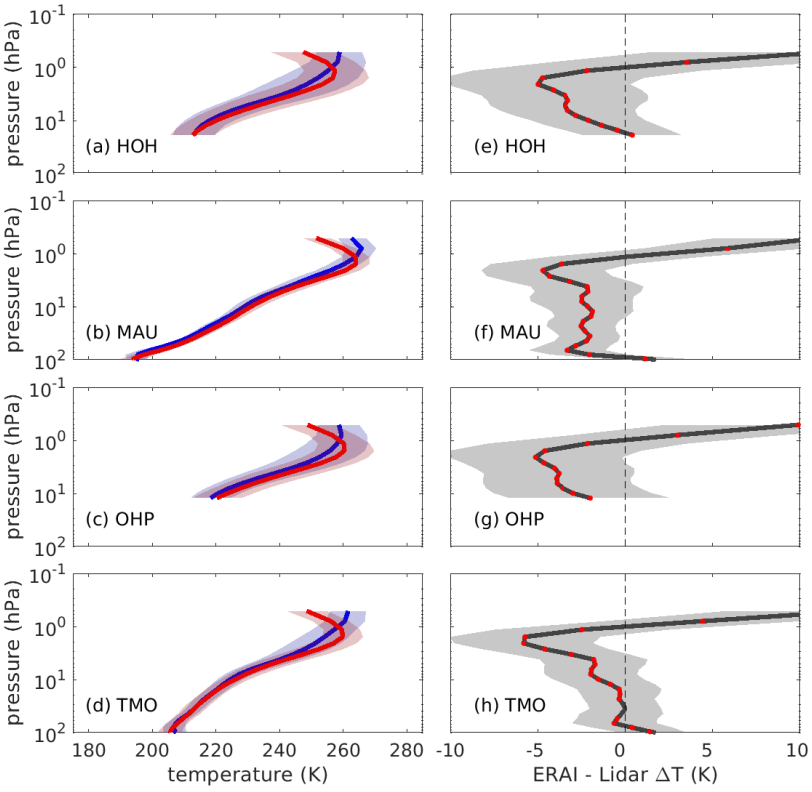

Figure 1. Mean profiles of matched temperatures from ERA-interim (blue) and Rayleigh temperature lidar (red) positioned at a) Hohenpeissenburg, b) Mauna Loa, c) Observatoire de Haute-Provence and d) Table Mountain Observatory for winter months (ONDJFM) between 1990 and 2017. Shading depicts 1 standard deviation in the mean temperature. The vertical profiles of the mean of the differences between ERA-interim and each lidar shown in a-d is shown in e-h respectively; shading shows 1 standard deviation of mean difference. Red dots show model levels where the mean of the differences were different from zero at the 95% significance level.

https://doi.org/10.5194/acp-2020-254
Preprint. Discussion started: 14 April 2020

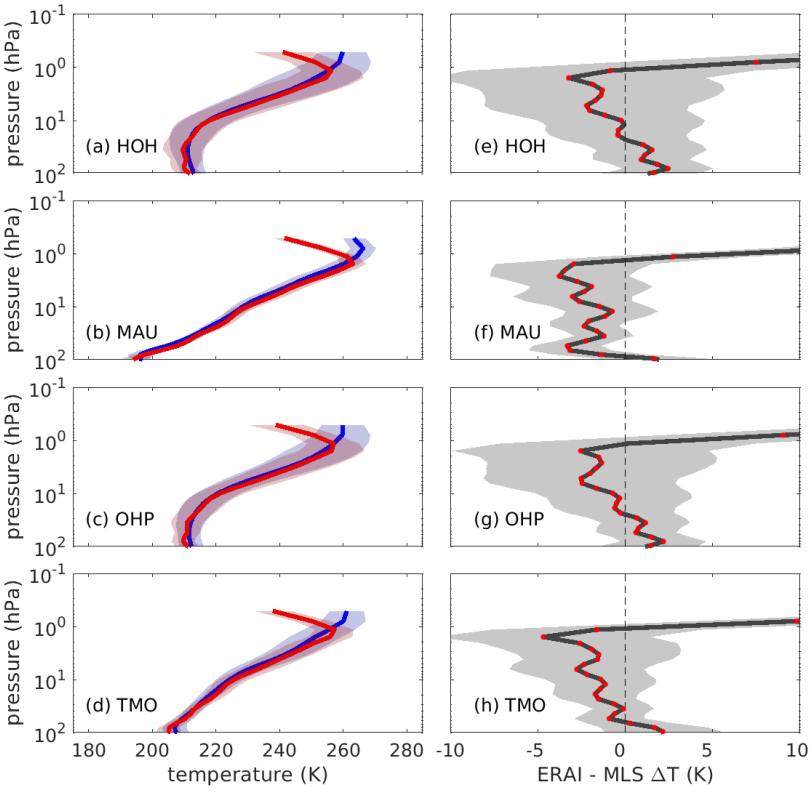

Figure 2. Mean profiles of matched temperatures from ERA-interim (blue) and EOS MLS passes (red)at a) Hohenpeissenburg, b) Mauna Loa, c) Observatoire de Haute-Provence and d) Table Mountain Observatory for winter months (ONDJFM) between 2000 and 2017. Shading depicts 1 standard deviation in the mean temperature. The vertical profiles of the mean of the differences between ERA-interim and EOS MLS shown in a-d is shown in e-h respectively; shading shows 1 standard deviation of mean difference. Red dots show model levels where the mean of the differences was different from zero at the 95% significance level.

https://doi.org/10.5194/acp-2020-254
Preprint. Discussion started: 14 April 2020

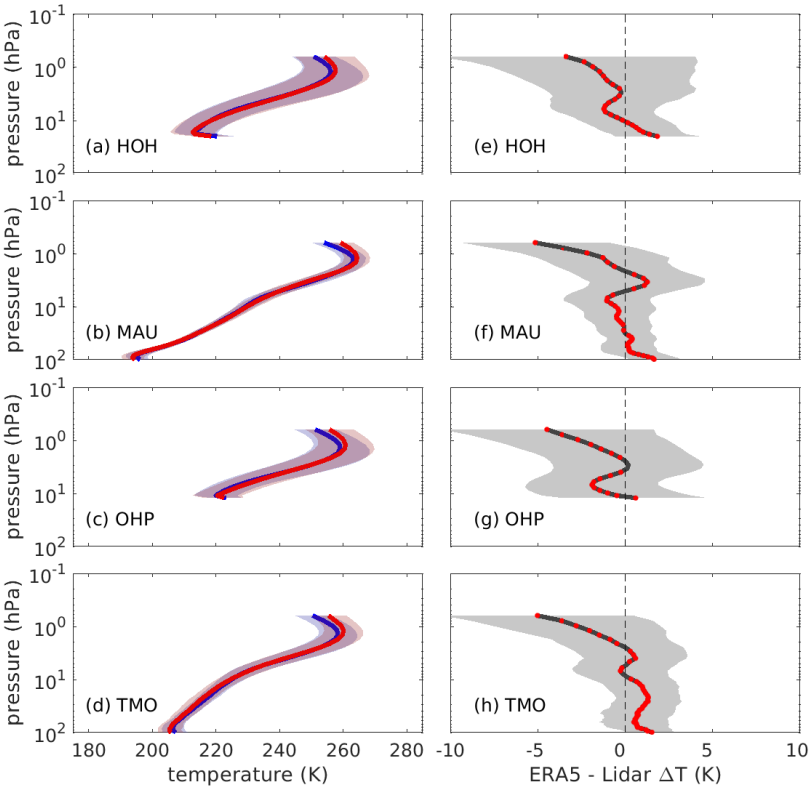

Figure 3. Mean profiles of matched temperatures from ERA-5 (blue) and Rayleigh temperature lidar (red) positioned at a) Hohenpeissenburg, b) Mauna Loa, c) Observatoire de Haute-Provence and d) Table Mountain Observatory for winter months (ONDJFM) between 1990 and 2017. Shading depicts 1 standard deviation in the mean temperature. The vertical profiles of the mean of the differences between ERA-5 and each lidar shown in a-d is shown in e-h respectively; shading shows 1 standard deviation of mean difference. Red dots show model levels where the mean of the differences were different from zero at the 95% significance level.

https://doi.org/10.5194/acp-2020-254
Preprint. Discussion started: 14 April 2020

Atmospheric Chemistry and Physics Discussions

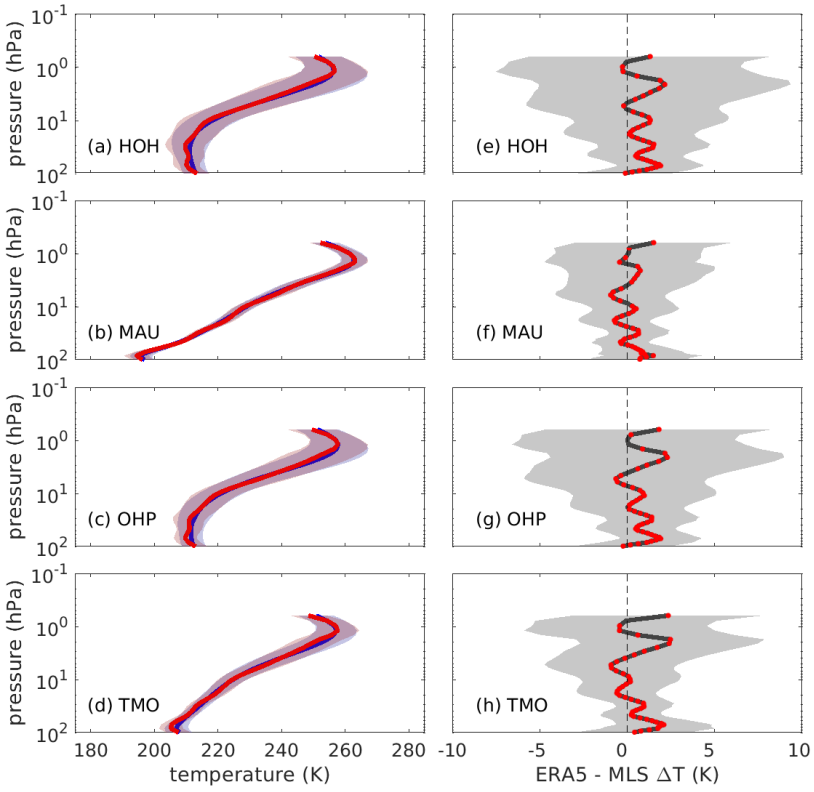

Figure 4. Mean profiles of matched temperatures from ERA-interim (blue) and EOS MLS passes (red)at a) Hohenpeissenburg, b) Mauna Loa, c) Observatoire de Haute-Provence and d) Table Mountain Observatory for winter months (ONDJFM) between 1990 and 2017. Shading depicts 1 standard deviation in the mean temperature. The vertical profiles of the mean of the differences between ERA-5 and EOS MLS are shown in a-d is shown in e-h respectively; shading shows 1 standard deviation of mean difference. Red dots show model levels where the mean of the differences was different from zero at the 95% significance level.

https://doi.org/10.5194/acp-2020-254
Preprint. Discussion started: 14 April 2020

Atmospheric Chemistry and Physics Discussions

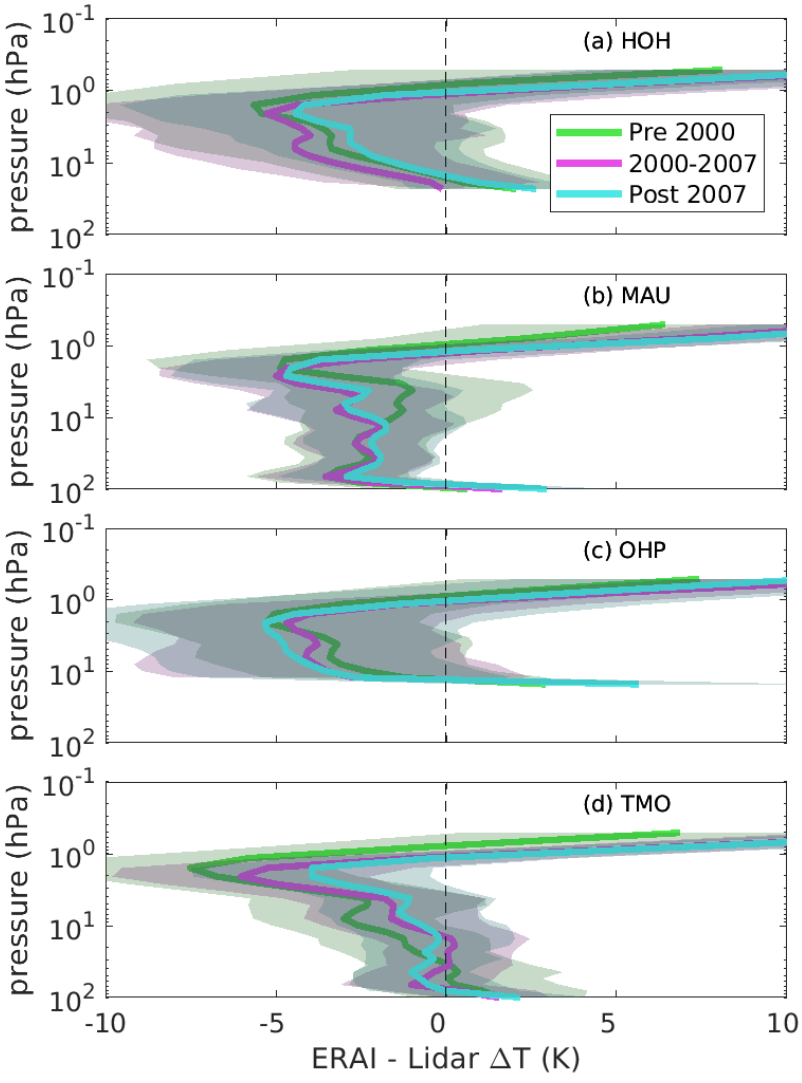

Figure 5. Temperature bias between ERA-interim and temperature lidars at (a) Hohenpeissenburg, (b) Mauna Loa, (c) Observatoire de Haute-Provence and (d) the table mountain observatory for the winter months ONDJFM. Green is ERA-interim comparisons from 1990 to 2000, purple is ERA-interim comparisons from 2000 to 2007 and blue is ERA-interim comparisons from 2007-2017. The solid lines are the mean and the shading is one standard deviation of the means.

https://doi.org/10.5194/acp-2020-254
Preprint. Discussion started: 14 April 2020

Atmospheric Chemistry and Physics Discussions

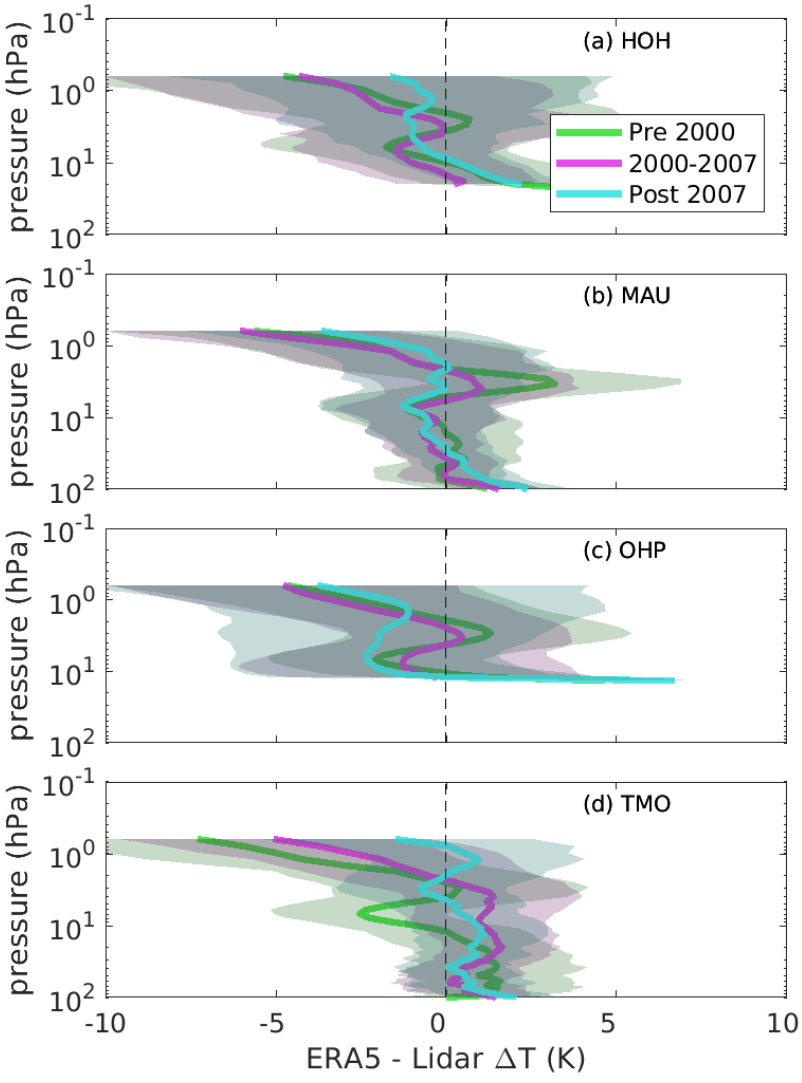

Figure 6. Temperature bias between ERA-5 and temperature lidars at (a) Hohenpeissenburg, (b) Mauna Loa, (c) Observatoire de Haute-Provence and (d) the table mountain observatory for the winter months ONDJFM. Green is ERA-5 comparisons from 1990 to 2000, purple is ERA-5 comparisons from 2000 to 2007 and blue is ERA-5 comparisons from 2007-2017. The solid lines are the mean and the shading is one standard deviation of the means.

https://doi.org/10.5194/acp-2020-254
Preprint. Discussion started: 14 April 2020

Atmospheric Chemistry and Physics Discussions

Table 1. Table summarising the geo-spatial and technical information of the 6 NDACC lidars used in this study

Lidar	Lat.^{o}	Lon.^{o}	Operating period[1]	Wavelength (nm)	Range gate Δz (m)
Hohenpeissenburg, Germany	47.8 N	11.0 E	1987-present	308	300
Mauna Loa, Haiwaii	19.8 N	155.7 W	1993 - present	355	300
Observatoire de Haute-Provence (OHP), France	43.9 N	5.7 E	1990-2016	532	1000
Table Mountain Observatory, US	34.5 N	117.7 W	1988-2014	352	300

[1] based on data availability