Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1029/2018JD028742

Validation of SABER v2.0 Operational Temperature Data With Ground-Based Lidars in the Mesosphere-Lower Thermosphere Region (75–105 km)

E. C. M. Dawkins[1,2], A. Feofilov[3], L. Rezac[4], A. A. Kutepov[1,2], D. Janches[1], J. Höffner[5], X. Chu[6], X. Lu[7], M. G. Mlynczak[8], and J. Russell III[9]

[1]NASA Goddard Space Flight Center, Greenbelt, MD, USA, [2]Department of Physics, Catholic University of America, Washington, DC, USA, [3]Laboratoire de Météorologie Dynamique, IPSL/CNRS, UMR8539, Ecole Polytechnique, Paris, France, [4]Department of Planets and Comets, Max-Planck-Institut für Sonnensystemforschung, Göttingen, Germany, [5]Leibniz-Institute for Atmospheric Physics, University of Rostock, Kühlungsborn, Germany, [6]Cooperative Institute of Research in Environmental Sciences & Department of Aerospace Engineering Sciences, University of Colorado Boulder, Boulder, CO, USA, [7]Department of Physics and Astronomy, Clemson University, Clemson, SC, USA, [8]NASA Langley Research Center, Langley, VA, USA, [9]Center for Atmospheric Science, Hampton University, Hampton, VA, USA

Key Points:
- The first comparative study of SABER version 2.0 operational temperature retrieval product with high-resolution ground-based lidar profiles
- A comparison of collocated multiyear seasonal average temperatures at nine locations (low, middle, and high latitudes) is presented
- An overall agreement for SABER and lidar data sets, but the presence of relative SABER warm bias during polar summer requires further investigation

Supporting Information:
- Supporting Information S1

Correspondence to:
E. C. M. Dawkins,
erin.dawkins@nasa.gov

Citation:
Dawkins, E. C. M., Feofilov, A., Rezac, L., Kutepov, A. A., Janches, D., Höffner, J., et al. (2018). Validation of SABER v2.0 operational temperature data with ground-based lidars in the mesosphere-lower thermosphere region (75–105 km). *Journal of Geophysical Research: Atmospheres*, *123*, 9916–9934. https://doi.org/10.1029/2018JD028742

Received 6 APR 2018
Accepted 3 AUG 2018
Accepted article online 17 AUG 2018
Published online 12 SEP 2018

Abstract The National Aeronautics and Space Administration (NASA) Thermosphere Ionosphere Mesosphere Energetics and Dynamics (TIMED) Sounding of the Atmosphere using Broadband Radiometry (SABER) instrument performs near-global measurements of the vertical kinetic temperature (T_k) profiles and volume mixing ratios of various trace species (including O_3, CO_2, and H_2O), with data available from 2002 to present. In this work, the first comparative study of the latest publically available SABER version 2.0 operational retrieval is reported in order to assess the performance of satellite T_k profiles relative to high-resolution ground-based lidar profiles. Collocated multiyear seasonal average T_k profiles were compared at nine different locations, representing a variety of different latitudes. In general, the SABER v2.0 and lidar mean seasonal T_k profiles agree well, with the smallest absolute values of ΔT_k (z) (SABER minus lidar) found between 85 and 95 km, where the respective SABER and lidar uncertainties were smallest. At altitudes ≥100 km, the SABER T_k (z) typically exhibited warmer temperatures relative to the lidar T_k (z) profiles, whereas for altitudes ≤85 km, SABER T_k (z) was cooler. Relative to lidar, SABER tends to exhibit a warm bias during high-latitude summertime, with the reasons for this currently still unclear. Overall, SABER was able to reproduce the general latitude- and season-specific variations in the lidar T_k profiles and shown to be statistically similar for most seasons, at most locations, for the majority of altitudes, and with no overall bias.

1. Introduction

Radiative, chemical, and dynamic coupling exists between all parts of the atmosphere, and the mesosphere-lower thermosphere (MLT) region serves as a vital boundary region between the Earth's atmosphere and space; it is subject both to the effects of solar irradiation, the solar wind, and ionospheric impacts from above, in addition to the influence of upwards propagating atmospheric waves and dynamical forcing, and radiative pumping from the lower atmosphere (e.g., see Beig, 2011; Kutepov et al., 2006; Li et al., 2016; Plane, 2003, Plane et al., 2015). As such, the MLT region is very sensitive to perturbations, and temperature variability is easier to detect here than in the lower atmosphere where the diurnal temperature variation is much larger than the warming over the past 150 years (Plane et al., 2015). It is imperative to continue to accurately monitor these temperatures and maintain consistent and well-calibrated instrumental temperature records in order to better understand the nature of this temperature variability. In addition, an ability to reproduce the temperature response within this part of the atmosphere, at a variety of timescales, represents a key test of any model physics and chemistry. To date, models are not fully capturing characteristics relating to gravity wave activity, vertical, and interhemispheric coupling, solar cycle related impacts, and various atmospheric phenomena such as sudden stratospheric warming events, all of which impact the temperatures within the MLT region (e.g., see Forbes et al., 2014; Gan et al., 2017; Goldberg et al., 2012; Tan et al., 2012; Walterscheid & Christensen, 2016; Yamashita et al., 2010).

Ground-based lidar measurements are routinely made at locations across the globe (e.g., Chen et al., 2016; Fong et al., 2014; Gerding et al., 2016; She & Yuan, 2015; Yuan et al., 2014) and provide long-term

Table 1
Summary of the Operational SABER Retrieval Uncertainties for a Single Profile (Reproduced From saber.gats-inc.com)

	70 km	80 km	90 km	100 km	110 km
T_k precision [K]	1.0	1.8	3.6	6.7	15.0
T_k systematic error [K]	1.5	1.4	4.0	5.0	25.0
T_k accuracy [K]	1.8	2.3	5.4	8.4	29.2

data series that are relatively free of biases associated with instrument drift (i.e., systematic errors which increase in magnitude, across the measurement period; e.g., Leblanc et al., 1998). Despite providing excellent vertical and temporal resolution, a key limitation of ground-based lidars are that they make point observations and are—for the most part—limited to land areas, typically within the northern hemisphere, and with largely no coverage of ocean areas. The advent of satellite-based temperature measurements has allowed temperature to be assessed on a near-global and near-continuous basis. Therefore, the ongoing comparison and validation of new satellite temperature products with available ground-based lidar data is vital.

In this work, we compare the latest publically available Sounding of the Atmosphere using Broadband Radiometry (SABER) temperature data with a variety of available ground-based lidar data, at different latitudes. Our goal is to highlight possible systematic deviations as a function of altitude, as well as quantifying the bias for further scientific use of the SABER v2.0 kinetic temperature (hereafter, T_k) data set. In section 2, we describe the National Aeronautics and Space Administration (NASA) Thermosphere Ionosphere Mesosphere Energetics and Dynamics (TIMED)/SABER instrument and the satellite temperature data sets. In section 3, we outline the available ground-based lidar station data used in this work and outline the rationale behind our comparison methodology. We present the results of the nine SABER-lidar temperature comparisons in section 4, followed by discussion of the results and summary (section 5) and conclusions (section 6).

2. The NASA TIMED/SABER Instrument

Launched on 7 December 2001, the NASA TIMED satellite orbits at a nominal altitude of 625 km, in a non-Sun-synchronous orbit with an orbital inclination of 74.1° and a mean orbit period of 97 min. One of the instruments onboard, SABER, consists of a 10-channel broadband limb-scanning infrared radiometer covering the spectral range 1.27–17 μm, performing global measurements which provide information on T_k, and volume mixing ratios (VMRs) of various trace species including O_3, CO_2, and H_2O. SABER also provides global measurements of atomic species O and H and provides over 30 radiative heating and cooling rates as standard data products to the community. The SABER instrument delivers near-continuous measurements and scans between the Earth's surface to 400 km, with a vertical instantaneous field-of-view of approximately 2.0 km at 50 km altitude, and with vertical sampling of ~0.4 km. Further information on the SABER instrument and TIMED mission can be found in Mlynczak (1997) and Russell et al. (1999).

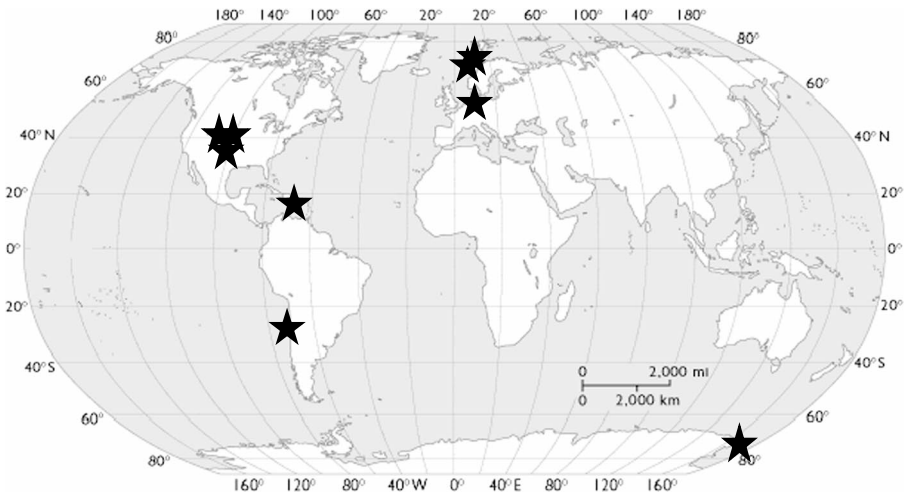

Figure 1. Map of global locations of ground-based lidar stations used in this study, which make temperature measurements of the mesosphere/lower thermosphere region.

Table 2
Summary of Lidar Measurements Used in This Study and Corresponding SABER Overflights

Lidar	Type	Lat	Lon	Approx. alt. range [km]	Data period used	Approx. vert. res. [km]	Uniform preprocessing	N_{SABER} overflights		
								Total available within range	Lidar in operation	% collocated
Spitsbergen (SPI)	K	78°N	15°N	70–115	2002–2003	0.25	Multi-year mean daily mean	2,261	81 days	N/A
ALOMAR (ALO)	Na	69.3°N	16.1°E	74–115	2002–2013	Varies between 0.15 and 0.87	Hourly mean	12,446	70	0.6%
Kühlungsborn, (KUH)	K	54.1°N	11.8°E	65–140	2002–2012	0.25	Multiyear mean daily mean	28,727	365 days	N/A
Boulder (BLD)	Na	40.1°N	105.2°W	74–130	2011–2014	0.96	Hourly mean	12,286	411	3.3%
Fort Collins (FTC)	Na	40.6°N	105°W	75–115	2002–2010	1.0	N/A	27,569	589	2.1%
Logan (LOG)	Na	41.7°N	111.8°W	75–115	2010–2015	0.85	Hourly mean	19,652	986	5.0%
Arecibo (ARE)	K	18.4°N	293.2°E	75–115	2003–2016	0.15	30-min intervals	40,128	178	0.4%
Cerro Pachon (AND)	Na	30.3°S	70.7°W	74–115	2010–2016	0.5	N/A	19,094	168	0.9%
McMurdo (MCM)	Fe	77.8°S	166.7°E	30–110	2011–2014	1.0	Hourly mean	5,154	285	5.5%

Note. The three-letter lidar station abbreviations used throughout the rest of this study are provided in the first column. The last three columns represent the number of available SABER overflights, including the total available within the vicinity of each lidar location (within ±5° latitude, ±15° longitude), the total available when the respective lidar was in operation, and the percentage of all available overflights spatiotemporally collocated with the lidar. N.B. The data for both SPI and KUH were available in a different format (as multiyear mean daily temperature means, only) to those of the other lidar locations, and thus, the number of days in which both lidar data and SABER overflights were available are stated here. Further information on these data sets can be found in the supporting information.

In this study, we use the latest publically available operationally retrieved SABER data, version 2.0 (hereafter, "v2.0"). The operational retrieval algorithm uses a combination of the measured 15-μm CO_2 (narrow channel) vertical emission profile and the diurnally averaged CO_2 vertical mixing ratio as calculated by the National Center for Atmospheric Research (NCAR) Whole Atmosphere Community Climate Model (WACCM). In 2015, a second SABER T_k data product became available, applying a 2-channel retrieval algorithm to obtain CO_2 VMR as well as T_k profiles in the 65- to 120-km altitude range (Rezac, Jian, et al., 2015). A preliminary validation by Rezac, Kutepov, et al. (2015) of the 2-channel retrieved T_k versus SABER operational (and Atmospheric Chemistry Experiment Fourier Transform Spectrometer) profiles revealed a general tendency of colder T_k profiles at higher latitudes (~5 K colder between 80 and 95 km in the summer polar region), although it is reported that individual profile-to-profile comparisons can show even larger deviations. Nevertheless, the 2-channel retrieval methodology can only be applied to daytime radiances, and therefore, in this work, we use only the single channel operational SABER T_k data set, which covers both day and night conditions, providing a higher chance for collocation with lidar data. More information on these two retrieval algorithms can be found on the instrument website: http://saber.gats-inc.com.

The $CO_2(\nu_2)$ populations related to the 15-μm emission are subject to departure from local thermodynamic equilibrium (LTE) conditions in the MLT (above ~65 km), making it necessary to include a vibrational non-LTE model into the retrieval algorithm. The rate at which the CO_2 molecule exchanges energy with other atmospheric molecules is controlled by collisional rate constants, which are not well known, and results in larger uncertainties being introduced into the derived T_k profiles. A more comprehensive description of the non-LTE processes influencing the 15-μm transitions are summarized in López-Puertas and Taylor (2001), and specifically relating to the SABER T_k retrieval in Mertens et al. (2001), Kutepov et al. (2006), García-Comas et al. (2008), Mertens et al. (2009), and Rezac, Jian, et al. (2015).

At 80 km, the total uncertainty in the v2.0 retrieved T_k due to both SABER random and systematic errors (e.g., pointing errors, instrument noise, pressure registration, and estimates for the CO_2 forward model) is approximately 5.5 K per single profile. At 90 km, this increases to 13 K and at 100 km constitutes 20.1 K. Table 1 presents a summary of the operational SABER retrieval uncertainties. A discussion of the previously available SABER product (v1.07) can be found in Remsberg et al. (2008) and García-Comas et al. (2008), along with updated available information on the instrument website at saber.gats-inc.com. Key updates to the v2.0 product since the v1.07 product include a recalibration of the SABER radiances and the use of retrieved [O] values

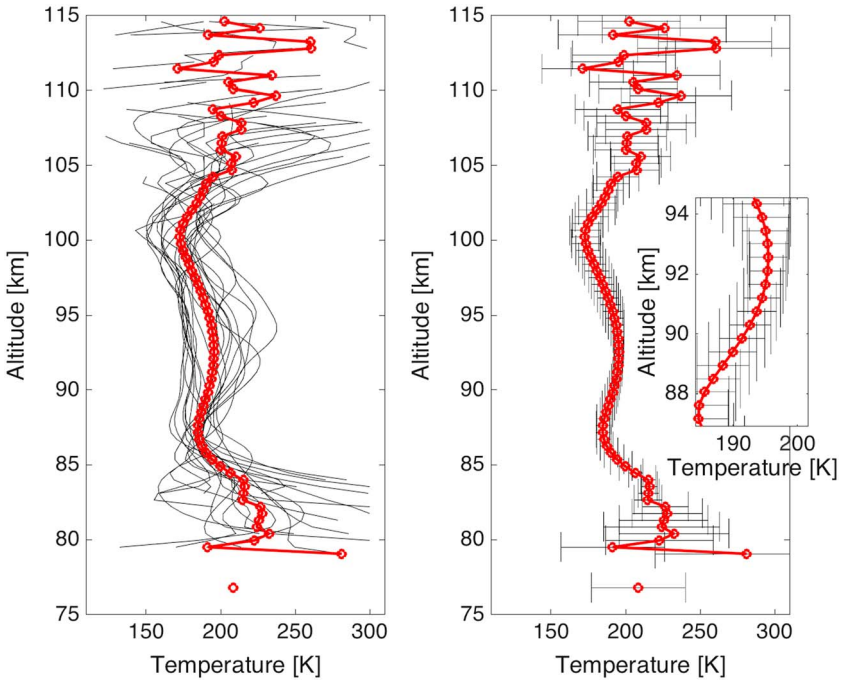

Figure 2. Daily vertical temperature profiles from Arecibo exhibit a large amount of variability. (left) All individual available lidar profiles for 19 December 2003 shown in black, daily mean shown in red. The approximate time between consecutive profiles is 30 min. (right) Mean daily temperature profile in red, with the mean associated lidar systematic and random measurement uncertainties shown as horizontal black bars.

for all data (in contrast to v1.07 which used retrieved [O] for daytime measurements only and where the solar zenith angle $< 85°$). Version 2.0 also uses CO_2 from an updated WACCM model, whereas the previous version used WACCM results that were scaled to match a CO_2 trend model. Finally, some reaction rates were changed in the CO_2 vibrational temperature model in v2.0.

3. Collocation With Available Ground-Based Lidar Temperature Data Sets and Methodology

For this study, nine lidar locations (visually presented in Figure 1 and described in Table 2) were selected in order to represent a range of different latitude positions. In addition, these lidar stations were chosen as they each either have relatively long-running data measurement series or a high availability of measurements, both of which make them well-suited for the nature of this comparison study. Nevertheless, these ground-based lidar stations rely on manual operation and weather conditions, and thus, the number of actual measurements performed can vary greatly between month-to-month and year-to-year; in many months, no measurements may be performed. Further information about each of the lidar data sets can be found in the supporting information.

A number of important contributing factors must be considered when making multi-instrument comparisons: (1) differences between the instruments and measurement methods, (2) the nature of the measured variable (in this case, the vertical temperature profile), and (3) issues related to sample size. First, the satellite and ground-based data sets constitute two very different types of measurements; SABER measures energy emitted from a volume as it scans the limb (with horizontal smearing of ~200–300 km) and involves the integration of an emitted radiance over time and latitude (e.g., see Rezac et al., 2011), while the lidar profile is a near-instantaneous vertical measurement, at a single location, with the temperature derived indirectly via observation of the meteoric metal layers as a tracer. The Doppler lidar system detects the Doppler broadening effects on the absorption spectra of metal atoms excited by multifrequency lidar light to infer temperatures, while the Boltzmann lidar system detects the population ratio

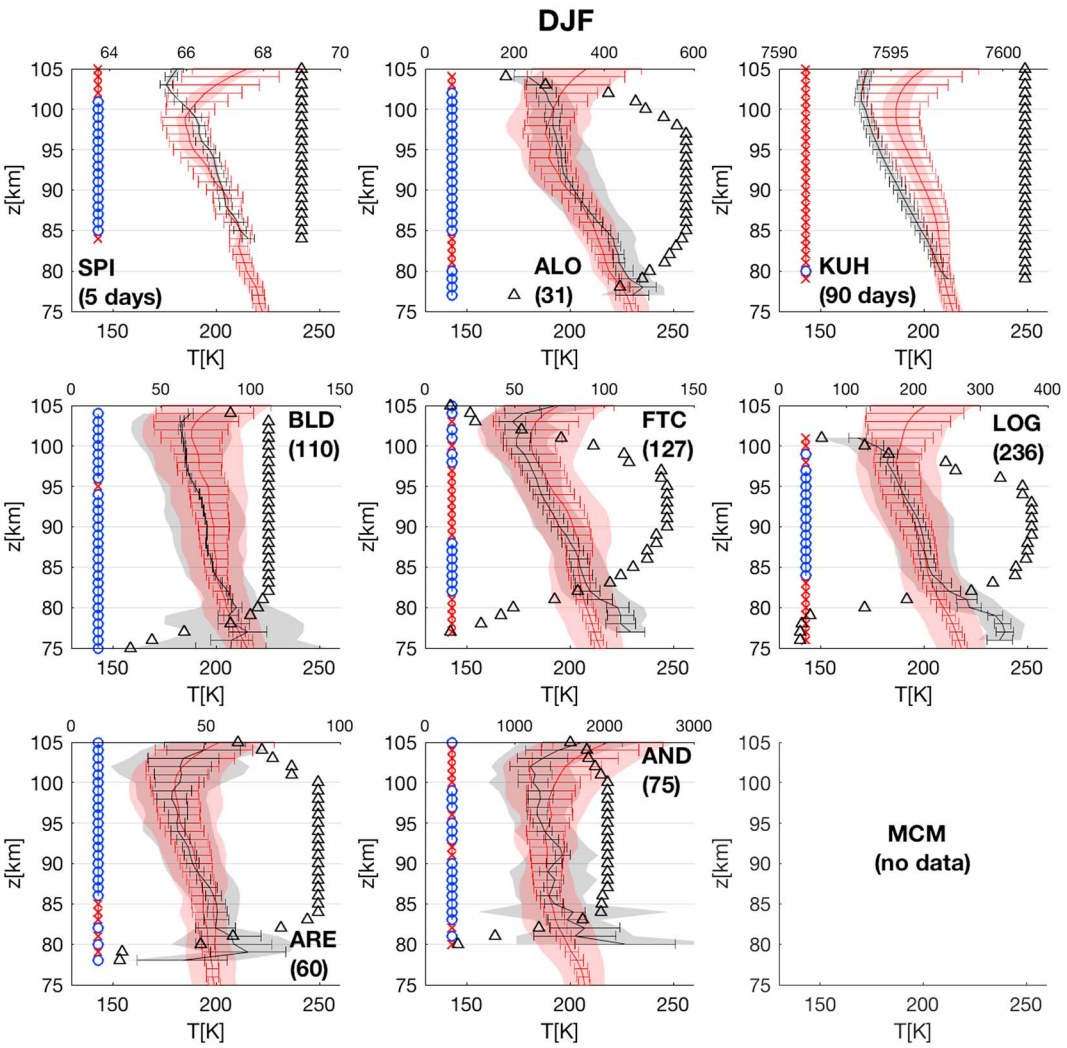

Figure 3. Comparison of multiyear December-January-February seasonal mean Sounding of the Atmosphere using Broadband Radiometry (SABER; red) and lidar (black) vertical temperature profiles, for each of the nine lidar station locations. The corresponding colored shaded regions represent ±1σ, and the horizontal bars represent the respective associated uncertainties. The Welch *t* test statistic between the SABER and lidar temperature data at each altitude is denoted on the left-hand side of each panel: The blue circle indicates that there is no significant difference between the SABER and lidar profile, while a red cross indicates that there is a significant difference ($p < 0.001$). The number of collocated SABER-lidar pairs is provided in brackets. The black triangles constitute the number of individual lidar profiles which contribute to the lidar seasonal mean at each altitude (see upper *x* axis).

between two different energy levels of Fe atoms via two lidar channels to infer temperatures (see a summary by Chu & Papen, 2005).

Second, strong temporal and spatial variability exists in the vertical temperature profile (with short-term variability primarily due to small-scale gravity waves and tides; e.g., see Chen et al., 2013; Dalin et al., 2013; Fritts & Alexander, 2003; Kutepov et al., 2007; Lu, Chen, et al., 2015; Lu, Chu, et al., 2015; Lu et al., 2017; Lübken et al., 2011; Mertens et al., 2004; Rapp et al., 2002; Stevens et al., 2017; Zhao et al., 2017): Unless the satellite and lidar instruments are perfectly collocated, there are likely to be substantial differences in the resulting temperature profiles as a result of this natural variability. Figure 2 illustrates the nature of the short-term variability in the vertical temperature profile, with all available nighttime profiles at the Arecibo lidar low-latitude location shown for 19 December 2003. Between 90 and 95 km, the derived temperatures can vary by up to 70 K. Outside of this altitude range, the lidar measurement and retrieval uncertainty increases significantly. Lastly, sample size is a vital consideration for comparison studies. Due to the strong

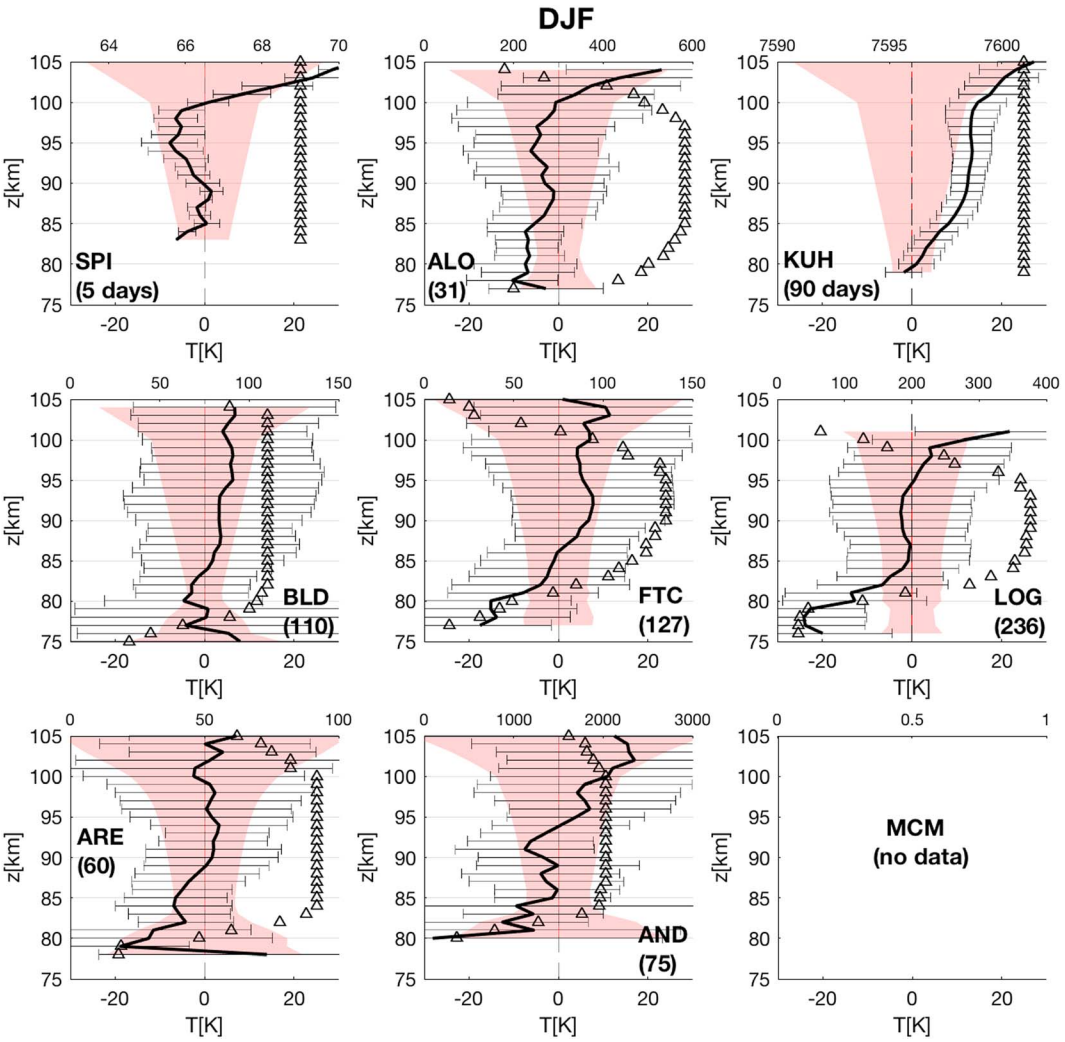

Figure 4. Comparison of differences between multiyear December-January-February seasonal vertical temperature profiles for each of the nine lidar station locations. For each panel, the solid black profile corresponds to the difference between the seasonal mean T_k profiles (SABER minus lidar). As before, the black triangles constitute the number of individual lidar profiles which contribute to the lidar seasonal mean at each altitude (see upper *x* axis). The red shaded region represents the altitude-dependent total mean SABER and lidar uncertainty (uncertainties added in quadrature). The exceptions to this are Spitsbergen and McMurdo (which do not have supplied lidar uncertainties), and thus, these red shaded areas consist of the SABER uncertainties only. The black horizontal bars represent the ±1σ of the differences between the T_k profiles for each of the individual collocated SABER-lidar pairs.

short-term variability in the T_k (*z*) profile, comparisons involving small sample sizes are likely to be nonrepresentative of the true performance of the satellite and data sets relative to one another and promote misleading conclusions.

In this work, a substantial effort has been made in order to minimize the effect of these differences (e.g., short-term variability produced by gravity wave activity and tides) by producing multiyear mean seasonal means. The spatiotemporal collocation criteria were carefully considered, and we follow a similar approach as that made by Remsberg et al. (2008), who used collocation criteria of ±5° latitude and longitude (centered around the location of each ground-based lidar station) and ±1 hr. In this work, however, we reduce the temporal coincidence criteria to ±30 min but expand the longitude-direction criterion to ±15° of the lidar station location (keeping the ±5° latitude-direction criterion the same). Further information on the lidar data sets and associated uncertainties and chosen methodology can be found in the

Table 3
Seasonal ΔT_k (z) Values at Each Location, for Selected Altitudes (Columns Corresponding to 75, 80, 85, 90, 95, 100, 105 km; ΔT_k (z) Defined as SABER-Lidar T_k (z))

DJF							
	75	80	85	90	95	100	105
SPI	-	-	+0.5	−0.4	−7.8	+0.8	+34.0
ALO	-	−7.4	−5.3	−3.0	−5.0	−0.6	-
KUH	-	+1.0	+8.2	+12.5	+13.3	+14.7	+27.2
BLD	+8.0	−4.7	+1.8	+3.3	+6.4	+5.1	-
FTC	-	−15.1	−1.0	+6.8	+5.5	+7.1	+0.9
LOG	-	−12.7	−0.8	−2.2	+0.6	+11.9	-
ARE	-	−12.5	−6.5	+1.7	+1.6	−2.4	+7.2
AND	-	−28.1	−1.3	−3.3	+3.9	+11.0	+12.5
MCM	-	-	-	-	-	-	-
MAM							
	75	80	85	90	95	100	105
SPI	-	-	−22.9	+3.1	+9.6	+23.8	+65.1
ALO	-	-	−9.7	−5.3	−7.5	-	-
KUH	-	−5.1	+12.7	+13.4	+16.4	+22.2	+47.4
BLD	-	+18.0	−0.1	−4.0	−8.7	−0.4	-
FTC	-	-	−0.0	+10.5	+8.8	−0.3	+11.7
LOG	-	−7.6	−1.2	+2.1	+0.3	-	-
ARE	-	−0.3	+2.5	−2.7	+0.7	+5.3	+14.2
AND	-	−29.6	−5.5	+1.3	+5.0	+3.0	+11.1
MCM	-	−6.1	−4.3	−0.2	−1.6	−1.5	+9.2
JJA							
	75	80	85	90	95	100	105
SPI	-	-	-	+15.0	+39.9	-	-
ALO	-	-	-	−5.9	+10.0	-	-
KUH	-	-	+8.7	+12.9	+21.0	+28.6	+55.1
BLD	-	+8.3	+0.5	+8.3	+12.2	+16.4	-
FTC	-	−8.1	−2.4	+2.3	+3.2	+7.0	+28.3
LOG	-	−10.1	−4.1	+1.4	+4.8	+13.1	-
ARE	-	−14.5	−21.2	−2.1	−4.1	+2.2	+9.3
AND	-	−22.0	−5.7	−8.8	−7.6	−2.1	+1.2
MCM	-	−14.9	−4.1	+1.0	+0.4	+5.1	+10.6
SON							
	75	80	85	90	95	100	105
SPI	-	-	-	-	-	-	-
ALO	-	-	−3.6	−14.0	−19.9	−10.1	-
KUH	-	+0.6	+10.3	+10.2	+9.8	+10.5	+25.2
BLD	-	+3.4	−0.8	+2.0	+3.8	+0.9	-
FTC	-	−15.4	−0.0	+1.1	−1.8	+0.3	+2.7
LOG	-	−9.7	−2.0	−4.0	−4.9	−0.7	-
ARE	-	−44.8	−17.9	−11.7	+6.4	+10.1	−11.2
AND	-	−21.7	−6.3	+0.7	−1.8	+1.8	+10.1
MCM	-	−10.6	+1.4	+4.2	+1.5	+3.8	-

supporting information. It should be noted that data from two of the lidar locations (Spitsbergen and Kühlungsborn) were available in a different averaging format to the other stations (preprocessed into multiyear daily averages) and thus required different processing, which is outlined further in the supporting information.

4. Comparison of Multiyear Mean Seasonal Temperatures: Results

4.1. December-January-February Comparisons

Figure 3 shows a comparison of the multiyear December-January-February (DJF) seasonal mean vertical temperature profiles for each of the nine lidar locations (N.B. no DJF data are available for MCM); the SABER profile (shown in red, with the shaded region representing the variability (±1σ), and the red bars representing the estimated single measurement uncertainty [systematic and random, discussed previously in section 2], all as a function of altitude) with the lidar data shown in black (shaded region corresponds to ±1σ, and the horizontal uncertainty bars represent the mean associated measurement uncertainty). The differences between the multiyear seasonal mean vertical SABER and lidar temperature profiles are shown later in Figure 4. The spatiotemporal collocation characteristics for all nine locations as a function of season are presented in Figures S1 and S2 in the supporting information.

Not all individual lidar T_k profiles extend across the full altitude range. The black triangles on the right-hand side of each panel correspond to the upper *x* axes and indicate the total number of individual lidar profiles included in the seasonal mean T_k as a function of altitude. The exceptions to this are for the SPI and KUH comparisons (where only preprocessed daily mean lidar *T* profiles were available), in which these triangles indicate the number of individual SABER overpass profiles included in the seasonal mean (N.B. see supporting information for more information on the differences between the lidar data sets and their processing).

The Welch's unequal variance *t* test statistic (Welch, 1947, 1951) is depicted on the left-hand side of each panel in Figure 3 and compares two independent samples with unequal variances (here SABER vs. lidar T_k (*z*)) and tests the hypothesis that the two populations have equal means. A significance value of *p*-value <0.001 (i.e., a 99% confidence level) is assigned. The blue circles indicate that the null hypothesis can be accepted (i.e., that the SABER and lidar mean profiles do not differ significantly to one another). A red cross indicates that the null hypothesis must be rejected and that the profiles differ significantly from one another.

Qualitatively, the SABER v2.0 and lidar T_k (*z*) profiles show similar overall vertical temperature structure to one another for nearly all locations, with each mean T_k (*z*) profile falling within the standard deviation and/or uncertainty range of the other data set. This is particularly true for the 85- to 100-km altitude region in which there is typically the most individual lidar measurements included in the overall seasonal mean T_k. Exceptions to this are sometimes on the top and bottom of the study altitude regions, in which the number of individual lidar measurements comprising the lidar seasonal mean T_k (*z*) decreases, for example, for LOG <85 km. At most locations, the 1σ ranges are similar for both satellite and lidar data sets for altitudes ≥80 km, which supports the assertion that the data sets are suitably collocated as both are capturing similar degrees of variability in the T_k (*z*) profiles. Quantitatively, the Welch *t* test

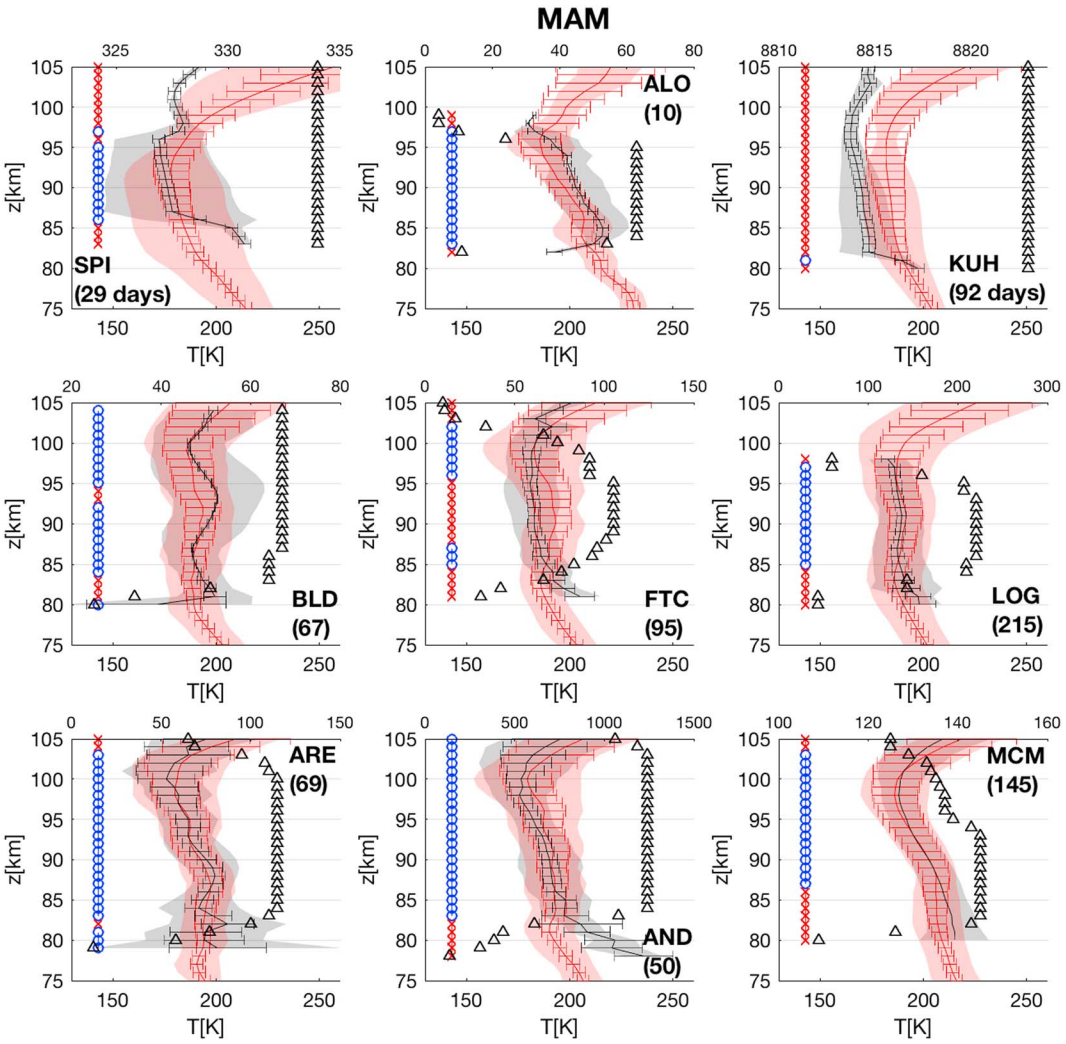

Figure 5. Comparison of multiyear March-April-May seasonal mean SABER (red) and lidar (black) vertical temperature profiles, for each of the nine lidar station locations. See Figure 3 for a description of each panel.

indicates no statistical difference between the SABER and lidar T_k profiles for the majority of compared altitudes at each location.

The differences between the multiyear seasonal mean vertical SABER and lidar T_k profiles are shown in Figure 4. For each panel, the combined uncertainty is shown (lidar and SABER uncertainties added in quadrature, for each altitude). The horizontal lines depict the $\pm 1\sigma$ of all of the individual differences between the collocated SABER-lidar temperature profiles. These can often be large as a result of combination of small-scale variability in the thermal profile, differences in spatiotemporal collocation, and differences in the viewing geometry of the SABER and lidar data sets.

For nearly all locations, and altitudes, the ΔT_k (z) curve falls within the combined uncertainty range, suggesting that the mean SABER and lidar T_k (z) profiles are similar enough to one another, within the uncertainty. The seasonal ΔT_k (z) values for each location at altitudes of 75, 80, 85, 90, 95, 100, and 105 km are reported in Table 3. The smallest ΔT_k (z) values generally occur between the 85- and 95-km altitude range. At 85 km and below, the SABER DJF T_k (z) is generally cooler than the corresponding lidar temperatures, while above 95 km, the SABER T_k (z) is generally warmer than the lidar data. The exception to this general pattern (SABER T_k (z) $>$ lidar T_k (z) above ~95 km and cooler than lidar T_k (z) below 85 km) is at KUH, where the SABER temperatures are consistently warmer than the lidar temperatures, ranging from +1.0 K at 80 km, +12.5 K at 90 km, +14.7 K at 100 km, to a maximum of +27.2 K at 105 km.

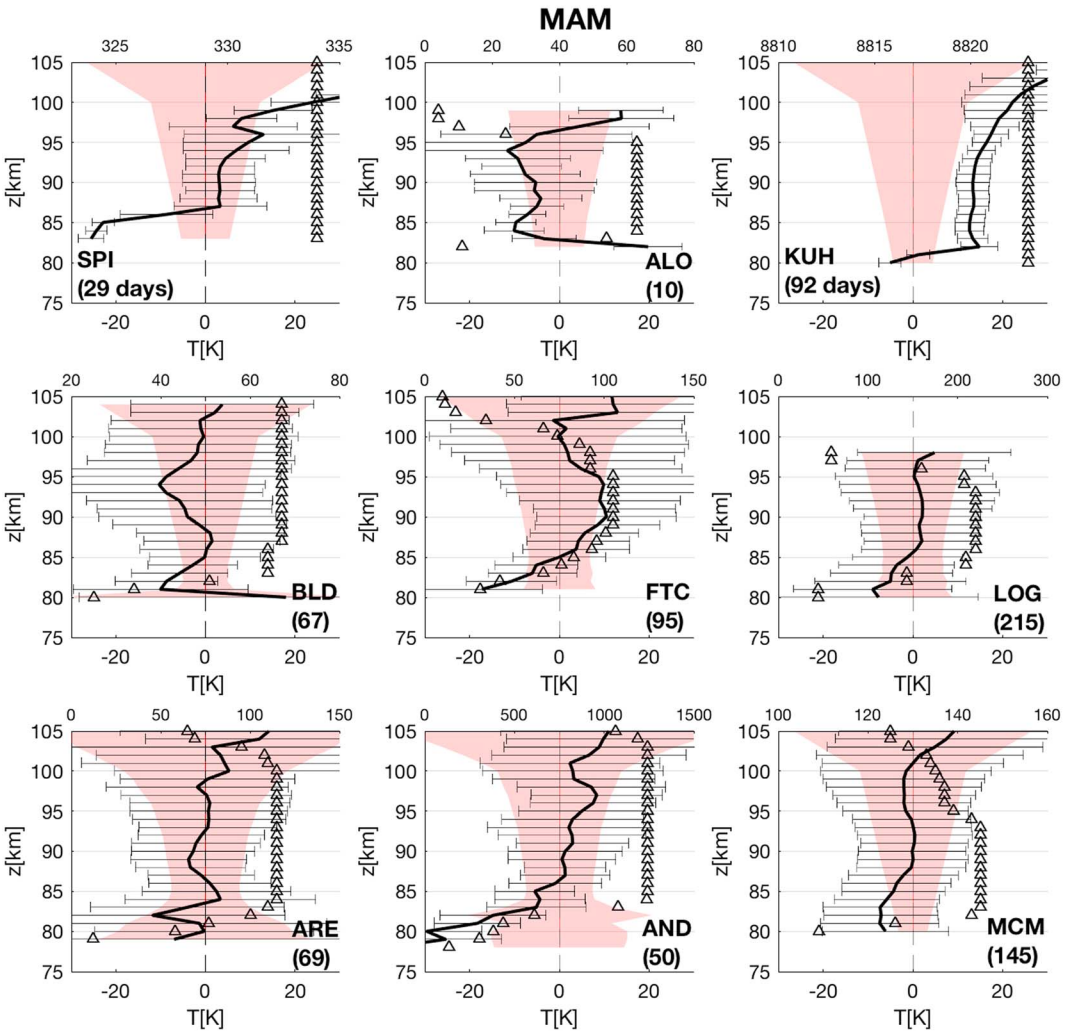

Figure 6. Comparison of differences between multiyear March-April-May seasonal vertical temperature profiles for each of the nine lidar station locations. See Figure 4 for a description of each panel.

4.2. March-April-May Comparisons

Similar to the DJF comparisons in section 4.1, Figure 5 shows the comparison of the multiyear seasonal SABER versus lidar T_k profiles for March-April-May (MAM), while Figure 6 shows the ΔT_k (z) between the SABER and lidar seasonal profiles for MAM. For nearly all locations, and nearly all altitudes, both the SABER and lidar seasonal MAM T_k (z) profiles and 1σ profiles exhibit a similar vertical structure to one another, with the Welch t test statistic indicating statistical similarity. This is particularly true for the ALO, LOG, ARE, AND, BLD, and MCM comparisons. The Welch t test indicates statistical agreement between the SABER and lidar between ~85 and 95 km at SPI and between 85 and 87 km and 96–102 K for FTC. The Welch t test indicates that the SABER and lidar T_k (z) profiles are statistically different for the majority of altitudes at KUH.

For the majority of locations and altitudes, the ΔT_k (z) falls within the combined uncertainty. The key exception to this is for KUH. The MAM ΔT_k (z) values for selected altitudes are also shown in Table 3, and similar to the DJF comparisons, the smallest ΔT_k (z) values typically occur between 85 and 95 km. Above this altitude range, the SABER temperatures are typically warmer than the corresponding lidar, with the opposite true at altitudes below.

4.3. June-July-August Comparisons

Figures 7 and 8 show the comparison of the seasonal SABER and lidar T_k profiles and ΔT_k (z) for June-July-August (JJA), respectively. Similar to both DJF and MAM, there is generally good agreement

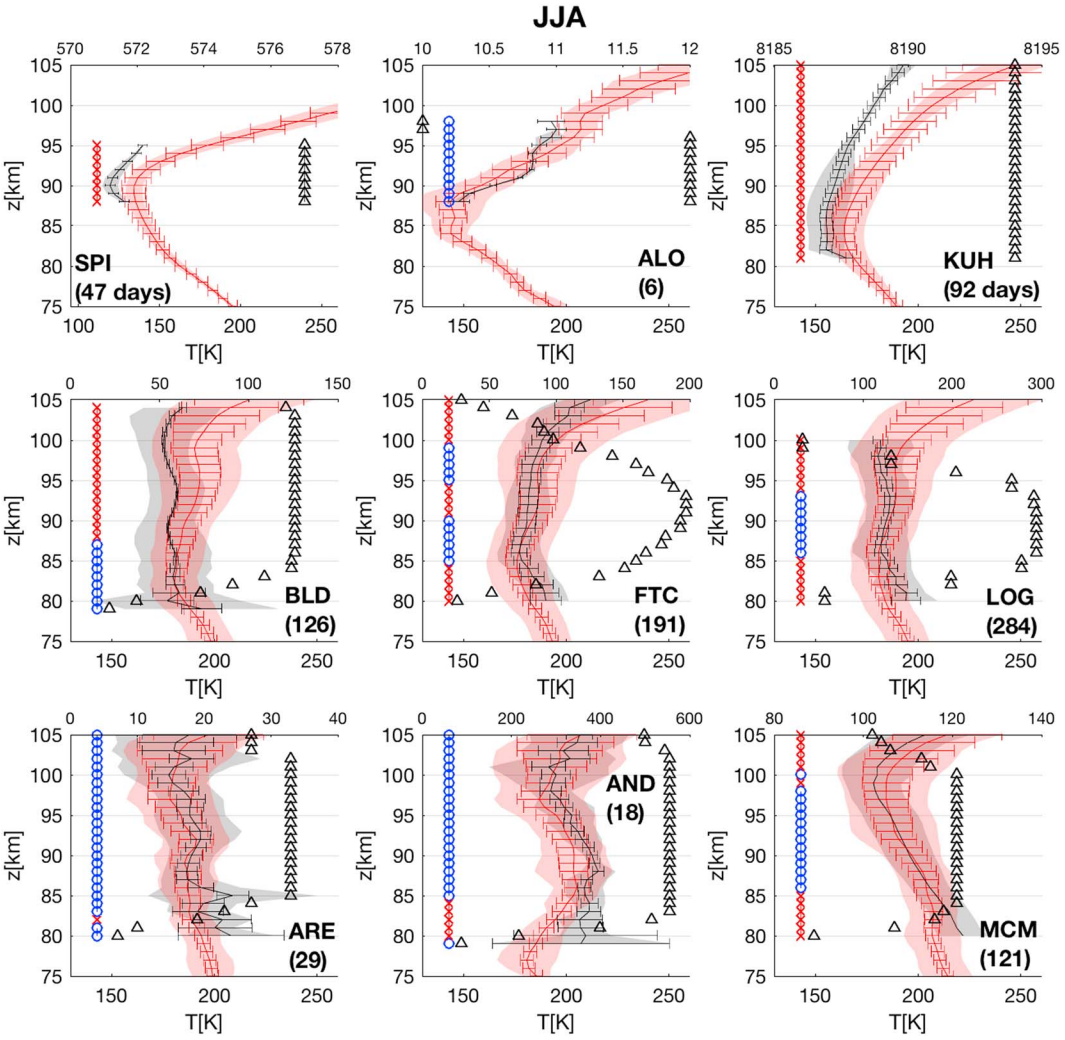

Figure 7. Comparison of multiyear June-July-August seasonal mean SABER (red) and lidar (black) vertical temperature profiles, for each of the nine lidar station locations.

between the SABER and lidar T_k (z) for most locations, with similar vertical T_k (z) profiles, similar 1σ ranges, and the Welch t test indicating statistical similarities, particularly for altitude ranges comprising of the largest number of individual lidar profiles. The exceptions to this are at SPI and KUH, where the Welch t test indicates that the respective SABER and lidar T_k (z) profiles are statistically different across the observed altitude range. Additionally, for altitudes ≥88 km, there is no statistical similarity between the SABER and lidar T_k (z) at BLD, although there is overlap between the 1σ and uncertainty ranges for each respective data set. The majority of the ΔT_k (z) values lie within the uncertainty ranges, with the main exceptions being the SPI and KUH profiles.

4.4. September-October-November Comparisons

Figure 9 shows the comparison of the SABER and lidar T_k (z) profiles for September-October-November (SON), while Figure 10 shows the corresponding ΔT_k (z) curves. No data exists for the SPI location, but the majority of the other locations indicate statistical similarity between the corresponding SABER and lidar T_k (z) profiles. The exceptions to this are for KUH and for the majority of altitudes for ALO and LOG. Despite this lack of statistical similarity, the SABER and lidar T_k (z) profiles typically show similar general vertical thermal structure to one another at these locations, and/or there is strong overlap between the 1σ and/or uncertainty ranges.

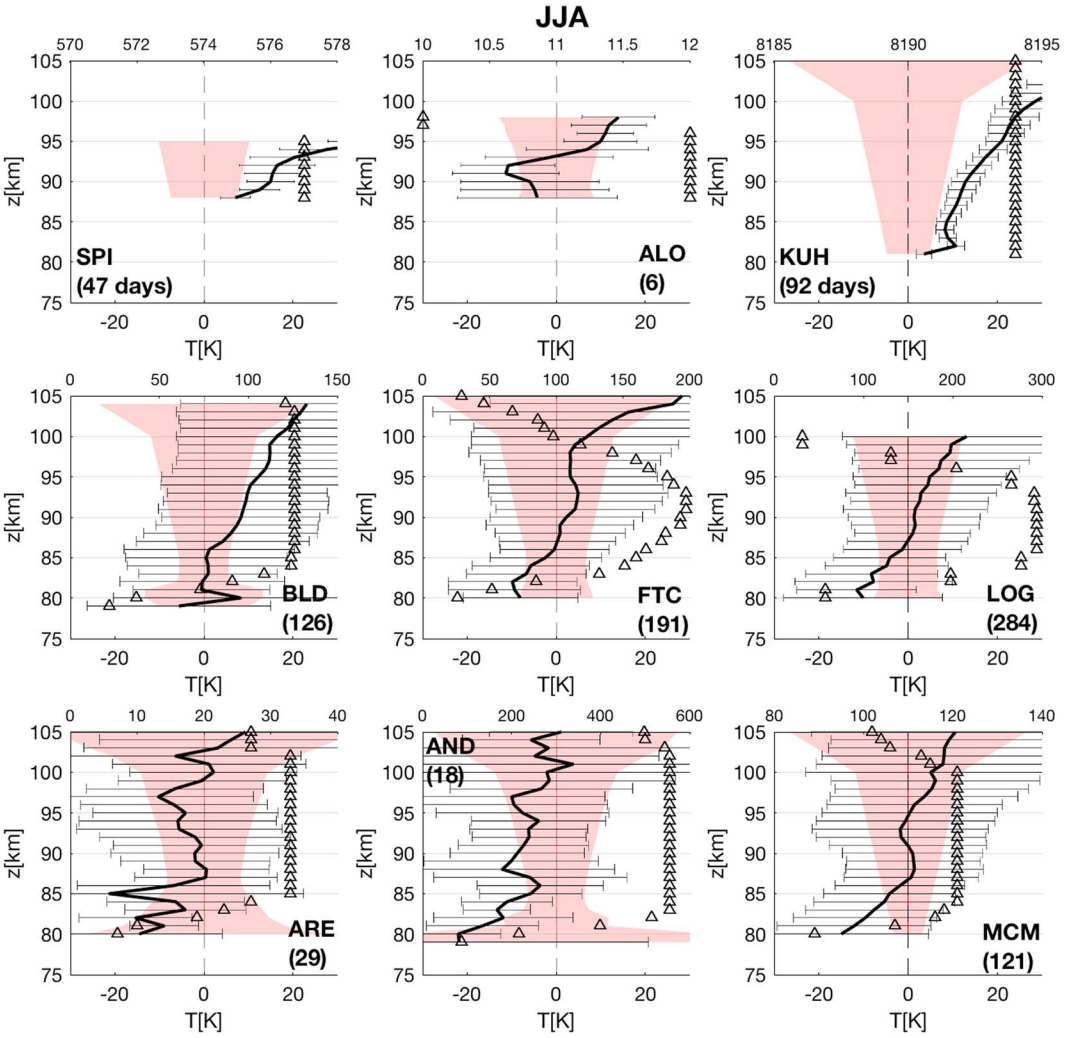

Figure 8. Comparison of differences between multiyear June-July-August seasonal vertical temperature profiles for each of the nine lidar station locations.

4.5. Overall Mean ΔT_k (z)

The mean ΔT_k (z) profiles for each location (averaged across all available seasonal mean profiles) are presented in Figure 11, along with the overall mean ΔT_k (z) profile for *all* locations and seasons. The overall mean ΔT_k (z) at each altitude (±1σ provided in brackets) is −9.9 (±9.7) K at 80 km, −3.1(±6.4) K at 85 km, +1.5(±5.8) K at 90 km, +3.6(±6.8) K at 95 km, +5.9(±6.8) K at 100 km, and +20.5(±18.8) K at 105 km (N.B. no overall mean is calculated at 75 km due to limited combined data points at this altitude). Excluding SPI and KUH, the overall mean ΔT_k (z) across all locations is −11.1(±9.7) K at 80 km, −3.8(±3.8) K at 85 km, −0.7(±4.1) K at 90 km, +0.4(±3.1) K at 95 km, +3.1(±4.2) K at 100 km, and +8.6(±2.6) K at 105 km.

4.6. Biases Between the SABER and Lidar Data Sets

To further explore the relationship between the SABER and lidar data sets, the individual collocated paired data are considered in Figure 12. For the majority of locations, the maximum of probability (filled red circles) lie along the 1:1 line indicating good overall agreement between the SABER and lidar collocated ΔT_k (z) pairs and no overall bias. The spread is likely due to natural atmospheric variability, spatio-temporal sampling inhomogeneity, and the difference in viewing geometries between the SABER and lidar observations. In contrast, there does seem to be a bias for temperatures below ~175 K for SPI with SABER generally observing warmer temperatures than the lidar. For temperatures >175 K, this apparent

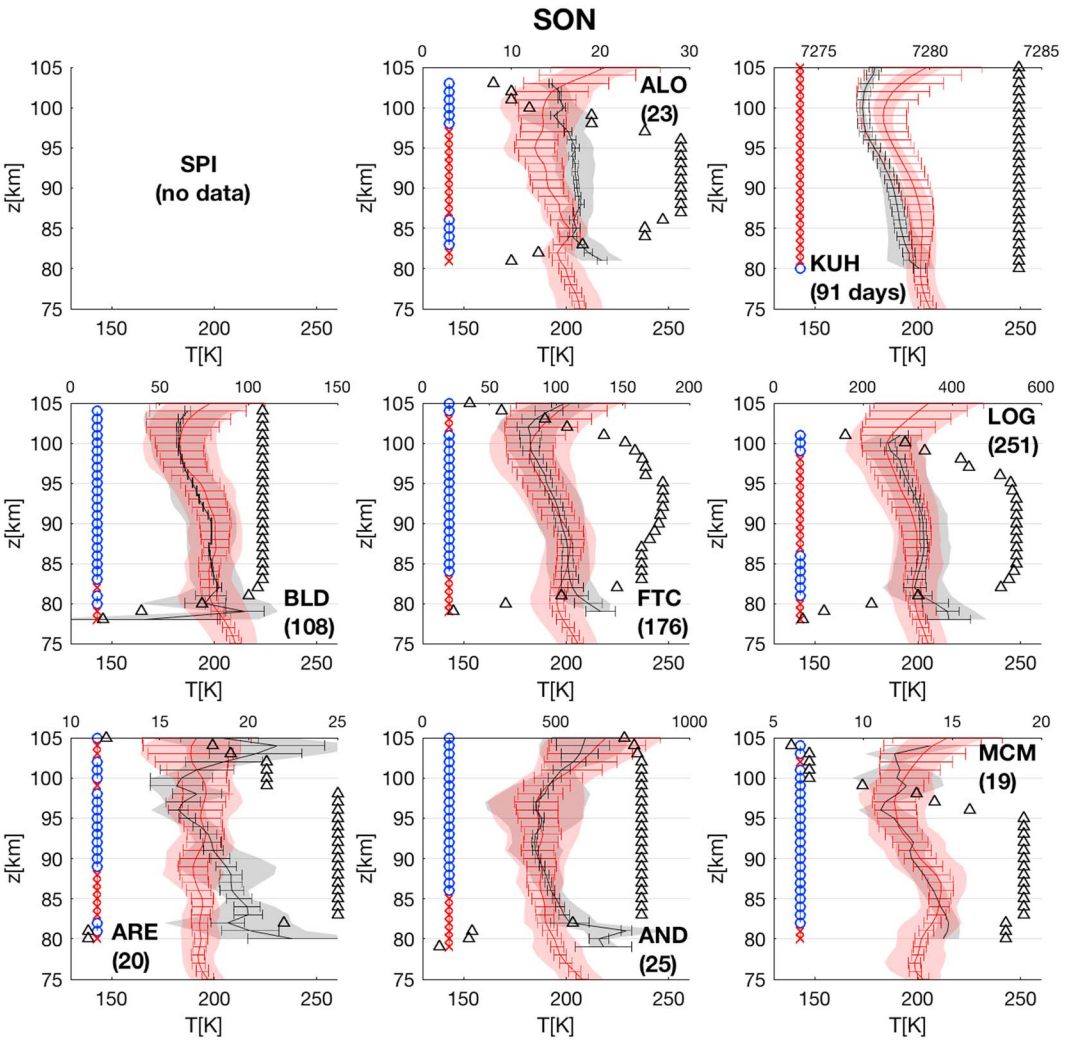

Figure 9. Comparison of multiyear September-October-November seasonal mean SABER (red) and lidar (black) vertical temperature profiles, for each of the nine lidar station locations.

bias is not present, with the maxima of probability lying closely along the 1:1 line. For KUH, there appear to be two separate distributions: one with the maxima of probability lying closely along the 1:1 line and the other in which SABER is consistently warmer. Additional analyses (not shown) have identified that these two populations have a seasonal basis; those lying closely along the 1:1 line typically correspond to days-of-year for DJF, MAM, and SON, with the second, consistently warmer population largely belonging to days-of-year associated with JJA.

Figure 13 shows the T_k (90 km) time series for both the SABER and lidar sets at each location, as a function of day-of-year. Also shown are the differences between the SABER and lidar means as a function of day-of-year. With the exception of KUH, there is uneven sampling per day-of-year for all data sets. Typically, the locations with the largest number of measurements (indicated by larger vertical standard deviation bars) have the smallest overall ΔT_k (90 km) across the full time period. Examples include FTC, where the overall ΔT_k (90 km) = +3.3 K, at LOG where the ΔT_k (90 km) is −1.2 K, at ARE where the ΔT_k (90 km) is −1.5 K, AND where the ΔT_k (90 km) is −1.8 K, and at MCM where the ΔT_k (90 km) is the smallest at −0.3 K. These results indicate that there is no overall bias between the SABER and lidar data sets for these locations. A specific exception to this is for the northern hemisphere summertime period for SPI and across the whole time period for KUH. In both cases, there is a clear bias in which the SABER temperatures are warmer than the collocated lidar data.

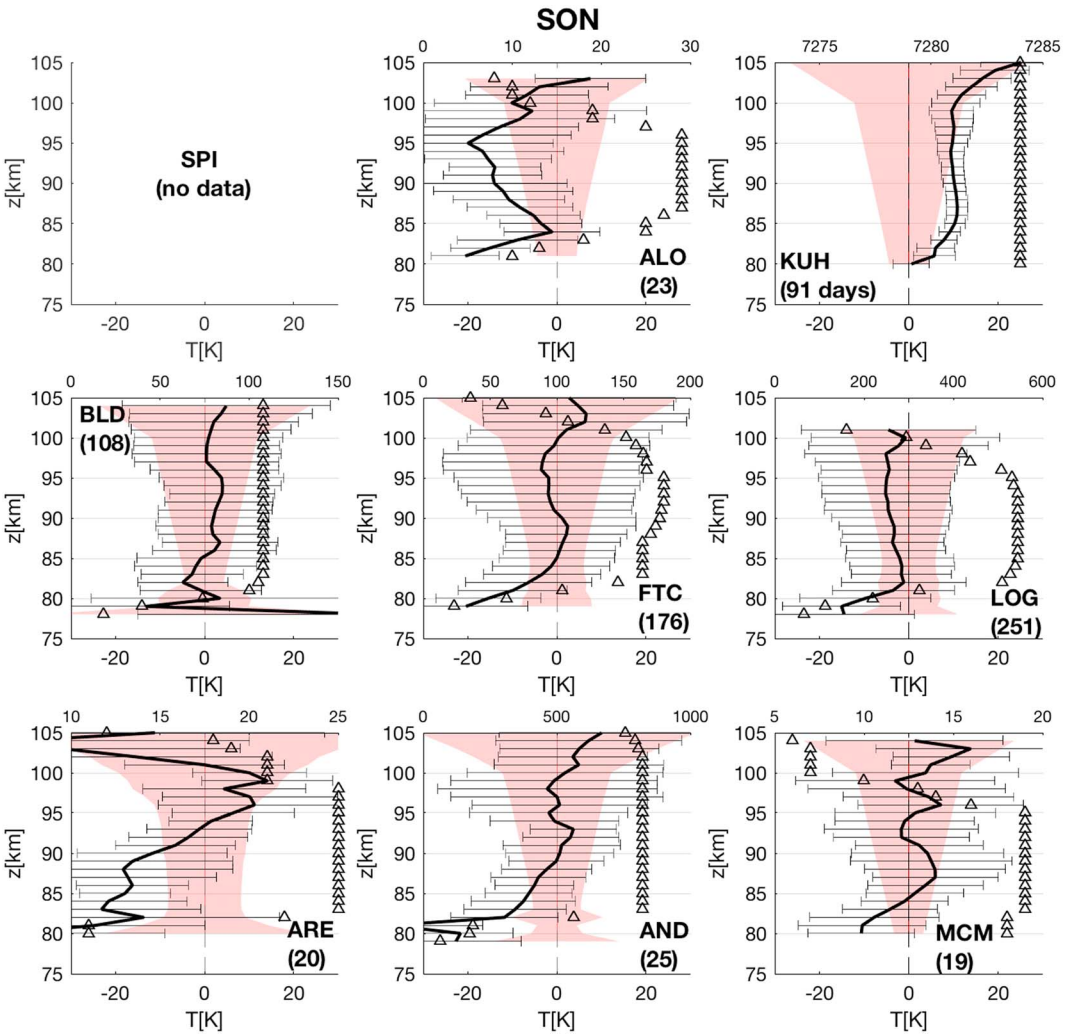

Figure 10. Comparison of differences between multiyear September-October-November seasonal vertical temperature profiles for each of the nine lidar station locations.

5. Discussion and Summary

Improved observations of the temperature variability within the neutral upper atmosphere can be used to better our understanding of this part of the atmosphere which is of considerable scientific interest. The ongoing development, comparison, and validation of new satellite temperature products are vital in creating and maintaining high-quality, near-continuous, and long-term observations of temperature within the MLT. Systematic comparisons with other temperature data sets will highlight issues which may require attention in further satellite retrieval algorithm versions.

We have presented the first comparative study of the collocated TIMED/SABER v2.0 multiyear seasonal mean profiles with those from nine different lidar station locations: Spitsbergen (78°N, 15°E), ALOMAR at Andøya (69°N, 16°E), Kühlungsborn (54°N, 12°E), Boulder (40°N, 105°W), Fort Collins (41°N, 105°W), Logan (42°N, 112°W), Arecibo (18°N, 293°E), Cerro Pachon (30°S, 71°W), and McMurdo (78°S, 167°E). We compared multiyear seasonal averages in an effort to assess the performance of the SABER v2.0 product relative to the high-resolution lidar temperature profiles. The observed SABER and lidar vertical temperature profiles are highly variable, particularly on short time scales principally as a result of gravity wave and tidal activity, and therefore, multiyear seasonal means were derived in order to mitigate and remove much of this short-term temperature variability signal in the data. An additional benefit to producing multiyear seasonal means is that it greatly enhances the collocated sample size, which makes the resulting means more representative of the seasonal vertical temperature profile.

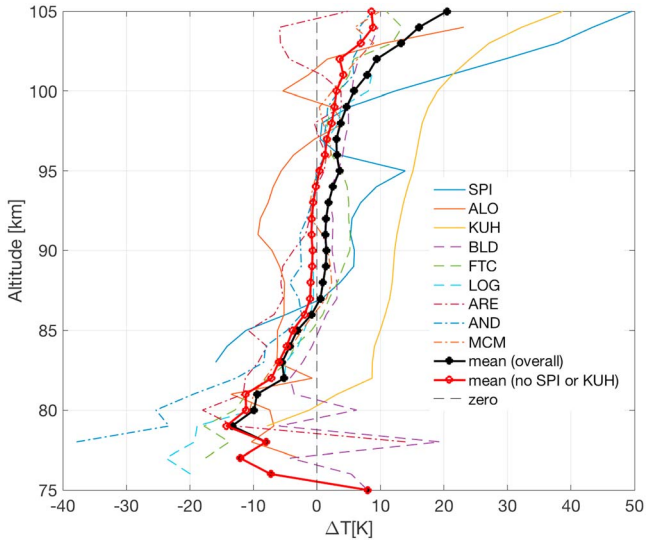

Figure 11. Mean ΔT_k (z) profile for each location and overall mean ΔT_k (z) for all locations.

Overall, the multiyear seasonally averaged T_k (z) derived from SABER are typically statistically similar to the corresponding lidar profiles for nearly all locations and for the majority of altitudes. Where there was good agreement between the mean SABER and lidar T_k (z) profiles, both data sets typically showed similar 1σ ranges, indicating that they were capturing comparable degrees of natural variability in the seasonal profiles. Additionally, the ΔT_k (z) curves generally fell within the combined uncertainties for both data sets, indicating satisfactory agreement between the collocated SABER v2.0 and lidar data sets where this was the case. For these locations, there was no obvious overall bias between the SABER and lidar data sets as shown in section 4.6.

For the comparison at Spitsbergen, there was good statistical agreement between the SABER and lidar data sets for DJF and MAM. This is also supported by the close agreement between the available data for the days-of-year corresponding to the DJF and MAM periods shown in Figure 13. Further, this relationship between these seasonally warmer temperatures (typically >175 K) can also be seen in Figure 12, where the maxima of probability for this temperature range generally lie along the 1:1 line. In contrast, there is a clear SABER warm bias relative to the lidar temperatures during the JJA time period (N.B. no data available for SON). The KUH panel in Figure 12 also exhibits similarity with the SPI panel; the maximum of probability indicates a SABER warm bias of ~10 K relative to lidar, with a distinct separate population indicating much warmer SABER temperatures. The KUH panel in Figure 12 also exhibits similarity with the SPI panel; the maxima of probability indicate a SABER warm bias of ~10 K relative to lidar, with a distinct separate population also indicating much warmer SABER temperatures. Additional analyses (not shown) identified this separate population primarily comprising of days-of-year corresponding to JJA. As such, it is not clear whether these differences are a consequence of the different comparison methodology at these locations, or deficiencies in the SABER retrieval during the cold summertime polar mesopause. Unlike the other locations, the SPI and KUH data sets are processed differently in that the lidar data is preprocessed into multiyear mean daily means, and the SABER data are collocated and averaged for each day. The SABER data include observations for all years within the multiyear period whereas for the same calendar day, the lidar data may only include a subset of years. As a result, the data sets are not as closely temporally collocated as for the other data sets, which likely contributes to the differences between the seasonal mean T_k (z) profiles, and may be biased by the effect of gravity waves (of various scales), tides, and planetary waves. Of particular influence may be larger scale gravity waves (with wavelengths of hundreds of kilometers, see Rapp et al., 2002) 2-day planetary waves with high zonal wave numbers (3 and 4; e.g., Meek et al., 1996; Muller & Nelson, 1978; Pogoreltsev, 1999; Rodgers & Prata, 1981), as well as effects due to large-scale mesospheric fronts/walls which produce significant temperature differences of the order 20–25 K across a small volume at the mesopause (Dalin et al., 2013).

However, another possibility is that the SABER operational retrieval produces profiles that are too warm during summer at high latitudes. Unfortunately, no hemispheric summertime (DJF) T_k is currently available for McMurdo, but limited summertime lidar T_k is available for Davis station, Antarctica (69°S, 78°E), as published in Lübken et al. (2015). A comparison of SABER with this published data set is presented in Figure S3; the mean SABER T_k profiles are warmer than the respective lidar mean T_k profiles, similar to both the SPI and KUH comparisons for JJA.

Independent observations, predominantly by falling sphere (FS) experiments, corroborate the respective lidar T_k profiles at SPI, KUH, and DAV (e.g., see Lübken et al., 1999; Lübken, 2000; Morris et al., 2012). Although the respective SABER summertime T_k profiles presented here lie within the T_k range of these independent data sets, the SABER mean is warmer than the lidar and FS data sets, often by ~10 K. This is consistent with previous work by Kutepov et al. (2006), who found that collocated polar summertime SABER v1.07 T_k data were ~10 K warmer than both a lidar climatology and a coincidental FS experiment for altitudes ≥87 km. In contrast to the other polar locations, the summertime comparison at ALOMAR (69°N) indicates good agreement between SABER and lidar T_k. FS and rocket grenade

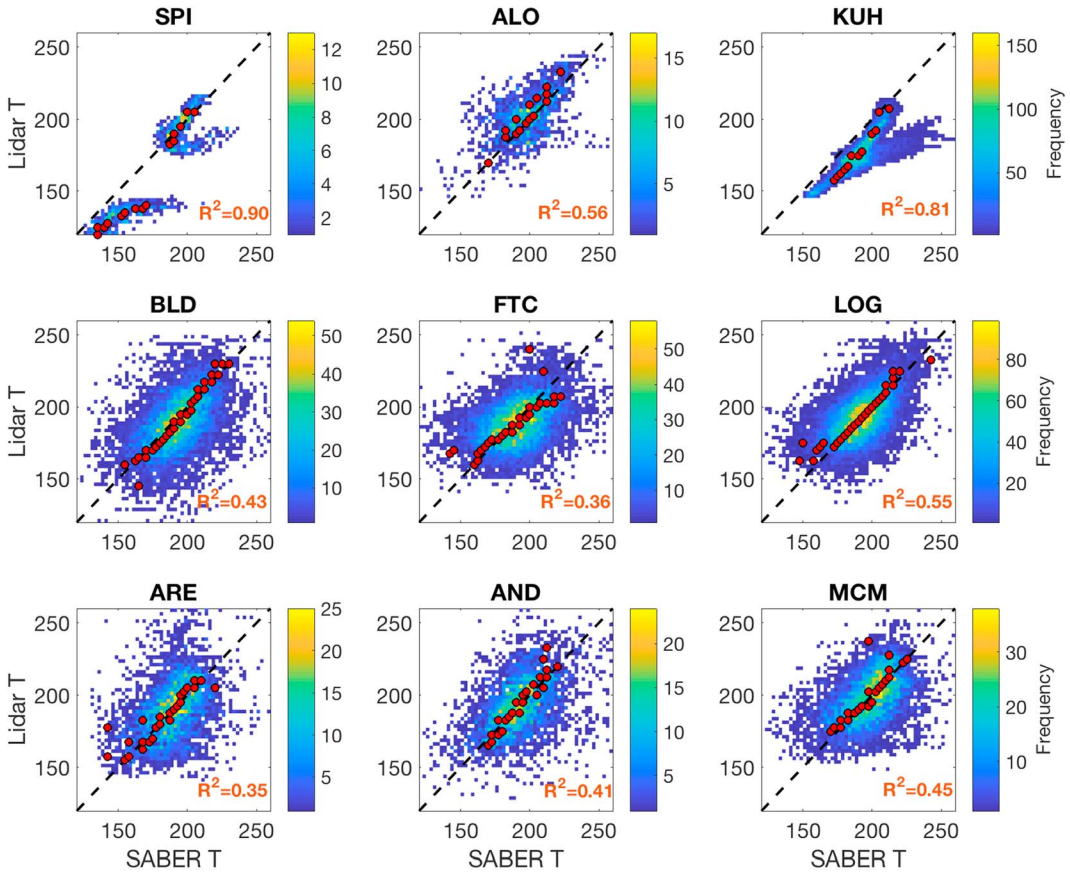

Figure 12. Correlation between individual pairs of collocated SABER and lidar T_k profiles, for all altitudes. The scatter points indicate the paired SABER and lidar T_k (z), and the color scale represents the number of occurrences within each 2.5 K temperature bin, as a function of both SABER and lidar T_k. The R^2 coefficient between the SABER and lidar T_k is stated in the lower right of each panel. The dashed black line represents the 1:1 line for a perfect correlation between SABER and lidar T_k. For each point along this 1:1, a Gaussian was fitted to all data points lying along a perpendicular transect. Where the data spread and statistics allowed a satisfactory fit, the maximal position of the Gaussian is plotted as a red filled circle. The red circles located above the 1:1 line indicate that the SABER T_k is cooler than the collocated lidar data. Where these red circles are located below the 1:1 line, the SABER T_k are warmer than the collocated lidar data.

experiments by Lübken (2000) at nearby locations (Point Barrow at 71°N and Kronogård at 66°N) exhibited similar temperatures and structure to the SABER and ALO mean T_k profiles presented in Figure 7, but both the SABER and ALO profiles represent the upper end of this range. Feng et al. (2017) reported a temperature offset of approximately +22 K, caused by issues with the laser sidebands in the laser system. Removing this offset would result in the ALO JJA T_k profile being more consistent with FS and rocket grenade observations (and in SABER showing a relative warm bias consistent with the other high-latitude summertime locations).

Across all seasons, the smallest location-specific ΔT_k (z) values occurred within the 85- to 95-km altitude range with the annual mean ΔT_k (z; shown in Figure 11) typically ranging between ±10 K, although for a number of locations, this was considerably smaller ranging between ±5 K or less. At altitudes ≥100 km, the SABER temperatures were often warmer than the lidar temperatures, ranging between approximately −5 K and +50 K. For altitudes ≤85 km, the reverse was true with SABER temperatures typically being cooler than those of the lidar profiles (range of approximately −38 K to +19 K). One simple explanation for the larger ΔT_k values at the top and bottom of the studied altitude range may lie in the limitations of the lidar data sets; the lidar can only measure in the presence of the metal layers (temperature is indirectly derived from the resonant scattering of light by these target species). The peak concentration of these metal layers is typically between 85 and 95 km, and outside of this altitude range, the signal-to-noise of the returned lidar signal is reduced, increasing the

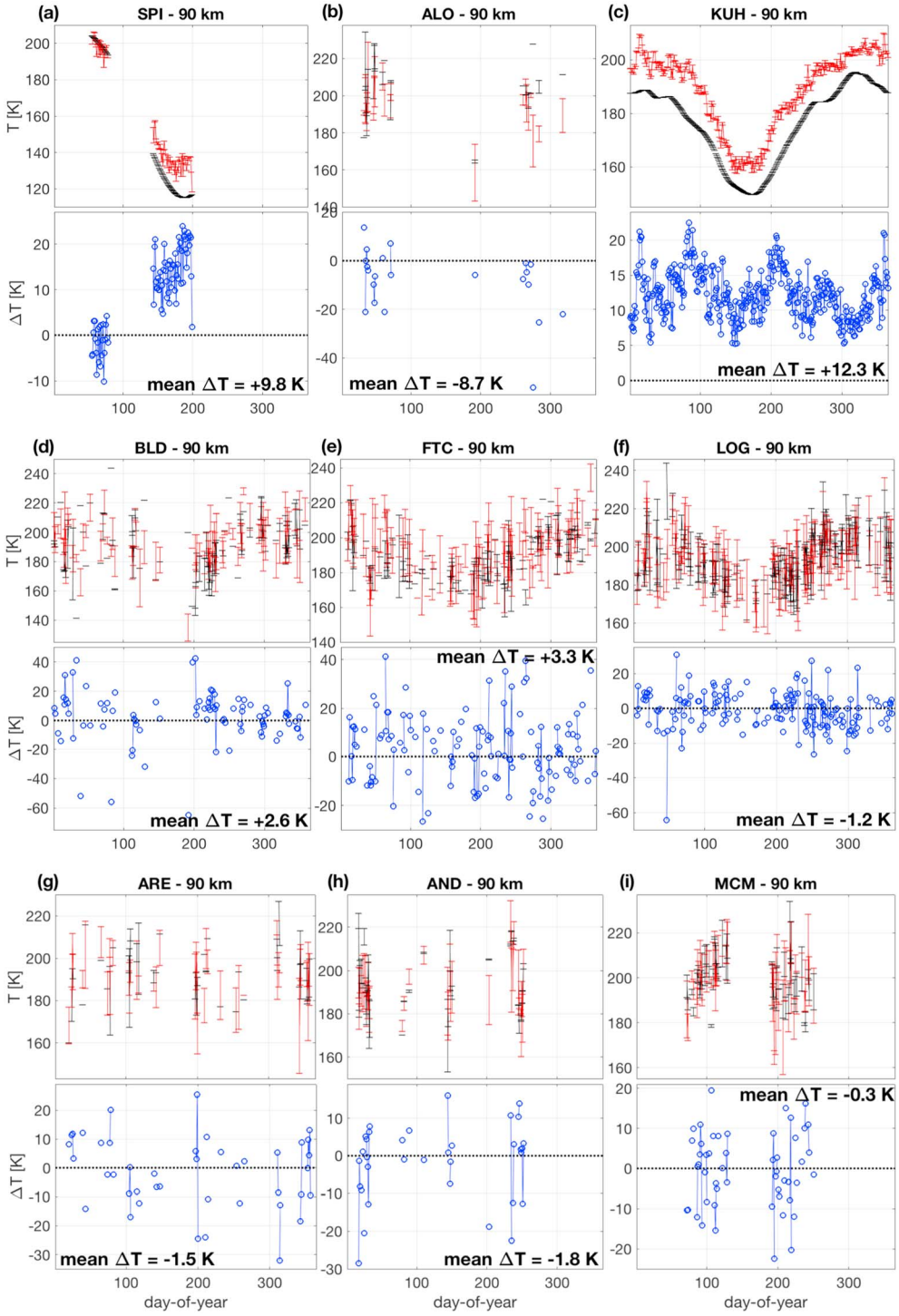

Figure 13. Time series plots of SABER and lidar data sets at an altitude of 90 km for each location: (a) Spitsbergen, (b) ALOMAR, (c) Kühlungsborn, (d) Boulder, (e) Fort Collins, (f) Logan, (g) Arecibo, (h) Cerro Pachon, and (i) McMurdo. The top panel in each location shows the multiyear mean T_k (90 km) and 1σ (SABER = red, lidar = black) as a function of day-of-year. The lower panel in each set shows the difference between the SABER-lidar mean T_k (90 km) as a function of day-of-year. The mean ΔT_k (90 km) value is stated in each case.

uncertainty. Additionally, these metal layers can be perturbed by dynamics and adiabatic motion. During the hemispheric summer these small-scale dynamics can compress these metal layers, which may cause both artifacts in the lidar data sets as well as producing a sharp vertical temperature gradient which the SABER instrument and retrieval may not be able to capture.

Another possible source of the larger ΔT_k (z) values may lie in the operational SABER T_k non-LTE model, which depends on several uncertain collisional rate coefficients and other inputs such as atomic oxygen. Remsberg et al. (2008) described the primary sources of systematic error relating to the forward model in the earlier version 1.07 SABER T_k retrieval, while García-Comas et al. (2008) investigated the propagation of non-LTE uncertainties into the retrieved v1.07 T_k. Remsberg et al. (2008) found that the uncertainties associated with the CO_2 VMR contributed one of the largest sources of systematic error in the non-LTE region, with uncertainty values of 1.3, 2.8, 3.6, 3.2, and 1.4 K for altitudes of 80, 85, 90, 95, and 100 km, respectively. Other key sources included the [O] profile, the associated physical quenching of $CO_2(\nu_2)$ with [O] (uncertainties associated with rate constant, $k_{\nu T}$; see Feofilov et al., 2012; Panka et al., 2017), and the uncertainties associated with $CO_2(\nu_2)$ vibrational quanta exchange rate constant ($k_{\nu\nu}$). García-Comas et al. (2008) demonstrated that the non-LTE region T_k (z) errors vary with both latitude and season and were largest during the polar summer (compared to a midlatitude summer location). At 92 km, the typical T_k (z) uncertainties associated with changes in the $CO_2(\nu_2)$-O quenching rate coefficient were ±1.3 K at the midlatitude location compared to ±5.0 K for the polar summer location, with the dependence being nonlinear with changes to [O] densities. At 96 km, these typical uncertainties increased to ±1.4 K (midlatitude) versus ±11 K (polar summer), ±3.6 K (midlatitude) versus ±18 K (polar summer) at 100 K, and to ±8.3 K (midlatitude) versus ±32 K (polar summer) at 104 km.

Although the sensitivity study by García-Comas et al. (2012) showed that neither reasonable CO_2 variations nor those of the non-LTE model parameters of the current models can explain biases higher than 5 K in this altitude region below 90 km, Feofilov et al. (2012) drew attention to the possible deficiency of current non-LTE models relating to the 15-μm CO_2 emission. These models currently do not account for possible nonthermal sources of CO_2 excitation by collisions with hot oxygen atoms and likely require additional study. Further work by Sharma (2015) suggested a possible mechanism of energy transfer of fast non-thermal O^3P atoms to the $CO_2(\nu_2)$ vibrations. Since current retrieval algorithms compensate missing excitation source by increasing retrieved temperature, this effect can be particularly important for summertime polar regions, where hot oxygen atoms are efficiently produced in the O_2/O_3 photolysis. Other additional notable sources of uncertainty include the horizontal T (p) gradients, which mainly affect high-latitude summer locations, and also determination of the reference pressure used in the retrieval (Remsberg et al., 2008).

6. Conclusions and Future Work

Overall, the comparisons show that the SABER v2.0 and lidar mean seasonal temperature profiles agree well, being statistically similar at the majority of altitudes, and showing similar thermal structure for most locations. Although these differences are largely within the uncertainties (and can thus be considered nonsignificant), SABER does show a warm bias relative to the lidar data sets at altitudes ≥100 km and a cool bias ≤85 km. SABER demonstrated an apparent summertime warm bias at SPI and KUH relative to the lidar data sets, which was supported by a supplementary comparison at DAV. The reasons for these differences are currently unsolved at this time and may be a combination both of differences in processing of the respective data sets and/or uncertainties in the retrieval that become more apparent at the summertime poles. Future work will investigate this further with a whole atmosphere model to investigate location-specific impacts on the SABER forward model and the resulting temperature profiles. Additionally, we plan to also include a comprehensive validation of the SABER "2-channel" using daytime lidar data sets where they exist.

Since 2002 the TIMED mission has, and continues to be, a fundamental source of measurements concerning the dynamics and energetics of the MLT. While still as productive now as when it was launched over 15 years ago, the planning of a follow-on mission is long overdue. Unfortunately, there are currently no planned missions to continue the rich data legacy of the SABER instrument. The likelihood of a resulting considerable data record gap represents a very real threat in the ability to detect trends due to differences in instrumental calibration, measurement technique, and algorithms. This pressing need for a successor to the SABER instrument has implications for MLT science, as well as the wider atmospheric community.

Acknowledgments

This work was supported by the TIMED project, NASA Heliophysics Division. L.R. was supported by the DFG priority program SPP1788 "Dynamic Earth." The work of A.F. was supported by the project "Towards a better interpretation of atmospheric phenomena" of the French National Program LEFE/INSU, managed by INSU-CNRS. The McMurdo lidar observations were supported by National Science Foundation (NSF) grants OPP-0839091, OPP-1246405, and OPP-1443726. The Boulder lidar observations were supported by NSF grants AGS-1136272 and AGS-1452351. Both X. C. and X. L. were partially supported by these NSF grants. The McMurdo lidar data were supplied by X. C. and X. L., while J. H. supplied the Spitsbergen and Kühlungsborn lidar data sets. The processed Spitsbergen, Kühlungsborn, and McMurdo data sets used within this manuscript can be found at doi: 10.5281/zenodo.1306736. All other lidar data were downloaded from the National Science Foundation CEDAR Madrigal website, http://cedar.openmadrigal.org, while the SABER operational version 2.0 data were obtained via http://saber.gats-inc.com.

References

Beig, G. (2011). Long-term trends in the temperature of the mesosphere/lower thermosphere region: 2. Solar response. *Journal of Geophysical Research*, *116*, A00H11. https://doi.org/10.1029/2011JA016646

Chen, C., Chu, X., McDonal, A. J., Vadas, S. L., Yu, Z., Fong, W., & Lu, X. (2013). Inertia-gravity waves in Antarctica: A case study with simultaneous lidar and radar measurements at McMurdo/Scott Base (77.8S, 166.7E). *Journal of Geophysical Research: Atmospheres*, *118*, 2794–2808. https://doi.org/10.1002/jgrd.50318

Chen, C., Chu, X., Zhao, J., Robers, B. R., Yu, Z., Fong, W., et al. (2016). Lidar observations of persistent gravity waves with periods of 3–10 h in the Antarctic middle and upper atmosphere at McMurdo (77.83°S, 166.67°E). *Journal of Geophysical Research: Space Physics*, *121*, 1483–1502. https://doi.org/10.1002/2015JA022127

Chu, X., & Papen, G. C. (2005). Resonance fluorescence lidar for measurements of the middle and upper atmosphere. In T. Fujii, & T. Fukuchi (Eds.), *Laser remore sensing* (pp. 179–432). Boca Raton, FL: CRC Press, Taylor & Francis Group.

Dalin, P., Connors, M., Schofield, I., Dubietis, A., Pertsev, N., Perminov, V., et al. (2013). First common volume ground-based and space measurements of the mesospheric front in noctilucent clouds. *Geophysical Research Letters*, *40*, 6399–6404. https://doi.org/10.1002/2013GL058553

Feng, W., Kaifler, B., Marsh, D. R., Höffner, J., Hoppe, U.-P., Williams, B. P., & Plane, J. M. C. (2017). Impacts of a sudden stratospheric warming on the mesospheric metal layers. *Journal of Atmospheric and Solar - Terrestrial Physics*, *162*, 162–171. https://doi.org/10.1016/j.jastp.2017.02.004

Feofilov, A. G., Kutepov, A. A., She, C.-Y., Smith, A. K., Pesnell, W. D., & Goldberg, R. A. (2012). CO2(v2)-O quenching rate coefficient derived from coincidental SABER/TIMED and Fort Collins lidar observations of the mesosphere and lower thermosphere. *Atmospheric Chemistry and Physics*, *12*(19), 9013–9023. https://doi.org/10.5194/acp-12-9013-2012

Fong, W., Lu, X., Chu, X., Fuller-Rowell, T. J., Yu, Z., Roberts, B. R., et al. (2014). Winter temperature tides from 30 to 110 km at McMurdo (77.8°S, 116.7°E), Antarctica: Lidar observations and comparisons with WAM. *Journal of Geophysical Research: Atmospheres*, *119*, 2846–2863. https://doi.org/10.1002/2013JD020784

Forbes, J. M., Zhang, X., & Marsh, D. R. (2014). Solar cycle dependence of middle atmosphere temperatures. *Journal of Geophysical Research: Atmospheres*, *119*, 9615–9625. https://doi.org/10.1002/2014JD021484

Fritts, D. C., & Alexander, M. J. (2003). Gravity wave dynamics and effects in the middle atmosphere. *Reviews of Geophysics*, *41*(1), 1003. https://doi.org/10.1029/2001RG000106

Gan, Q., Du, J., Fomichev, V. I., Ward, W. E., Beagley, S. R., Zhang, S., & Yue, J. (2017). Temperature responses to the 11-year solar cycle in the mesosphere from the 31 year (1979–2010) extended Canadian Middle Atmosphere Model simulations and a comparison with the 14 year (2002–2015) TIMED/SABER observations. *Journal of Geophysical Research: Space Physics*, *122*, 4801–4818. https://doi.org/10.1002/2016JA023564

García-Comas, M., Funke, B., López-Puertas, M., Bermejo-Pantaleón, D., Glatthor, N., von Clarmann, T., et al. (2012). On the quality of MIPAS kinetic temperature in the middle atmosphere. *Atmospheric Chemistry and Physics*, *12*(13), 6009–6039. https://doi.org/10.5194/acp-12-6009-2012

García-Comas, M., López-Puertas, M., Marshall, B. T., Wintersteiner, P. P., Funke, B., Bermejo-Pantaleón, D., et al. (2008). Errors in Sounding of the Atmosphere using Broadband Radiometry (SABER) kinetic temperature caused by non-local-thermodynamic-equilibrium model parameters. *Journal of Geophysical Research*, *113*, D24106. https://doi.org/10.1029/2008JD010105

Gerding, M., Baumgarten, K., Höffner, J., & Lübken, F. J. (2016). Lidar soundings between 30 and 100 km altitude during day and night for observation of temperatures, gravity waves and tides. In EPJ Web of Conferences (Vol. 119, p. 13001). *EDP sciences*. https://doi.org/10.5194/amt-9-3707-2016

Goldberg, R. A., Feofilov, A. G., Pesnell, W. D., & Kutepov, A. A. (2012). Inter-hemispheric coupling during northern polar summer periods of 2002–2010 using TIMED/SABER measurements. *Journal of Atmospheric and Solar - Terrestrial Physics*, *104*, 277–284. https://doi.org/10.1016/j.jastp.2012.11.018

Kutepov, A. A., Feofilov, A. G., Marshall, B. T., Gordley, L. L., Pesnell, W. D., Goldberg, R. A., & Russell, J. M. III (2006). SABER temperature observations in the summer polar mesosphere and lower thermosphere: Importance of accounting for the CO ν quanta V-V exchange. *Geophysical Research Letters*, *33*, L21809. https://doi.org/10.1029/2006GL026591

Kutepov, A. A., Feofilov, A. G., Medvedev, A. S., Pauldrach, A. W. A., & Hartogh, P. (2007). Small-scale temperature fluctuations associated with gravity waves cause additional radiative cooling of the mesopause region. *Geophysical Research Letters*, *34*, L24807. https://doi.org/10.1029/2007GL032392

Leblanc, T., McDermid, I. S., Keckhut, P., Hauchecorne, A., She, C. Y., & Krueger, D. A. (1998). Temperature climatology of the middle atmosphere from long-term lidar measurements at middle and low latitudes. *Journal of Geophysical Research*, *103*(D14), 17,191–17,204. https://doi.org/10.1029/98JD01347

Li, T., Calvo, N., Yue, J., Russell, J. M. III, Smith, A. K., Mlynczak, M. G., et al. (2016). Southern Hemisphere summer mesosphere responses to El Niño-Southern Oscillation. *Journal of Climate*, *29*(17), 6319–6328. https://doi.org/10.1175/JCLI-D-15-0816.1

López-Puertas, M., & Taylor, F. W. (2001). Non-LTE radiative transfer in the atmosphere. series on Atmospheric, Oceanic and Planetary Physics. *World Scientific*, *3*. https://doi.org/10.1142/4650

Lu, X., Chen, C., Huang, W., Smith, J. A., Chu, X., Yuan, T., et al. (2015). A coordinated study of 1-h mesoscale gravity waves propagating from Logan to Boulder with CRRL Na Doppler lidars and temperature mapper. *Journal of Geophysical Research: Atmospheres*, *120*, 10,006–10,021. https://doi.org/10.1002/2015JD023604

Lu, X., Chu, X., Fong, W., Chen, C., Yu, Z., Roberts, B. R., & McDonald, A. J. (2015). Vertical evolution of potential energy density and vertical wavenumber spectrum of Antarctic gravity waves from 35 to 105 km at McMurdo (77.8°S, 166.7°E). *Journal of Geophysical Research: Atmospheres*, *120*, 2719–2737. https://doi.org/10.1002/2014JD022751

Lu, X., Chu, X., Li, H., Chen, C., Smith, J. A., & Vadas, S. L. (2017). Statistical characterization of high-to-medium frequency mesoscale gravity waves by lidar-measured vertical winds and temperatures in the MLT. *Journal of Atmospheric and Solar - Terrestrial Physics*, *162*, 3–15. https://doi.org/10.1016/j.jastp.2016.10.009

Lübken, F.-J. (2000). Nearly zero temperature trend in the polar summer mesosphere. *Geophysical Research Letters*, *27*(21), 3603–3606. https://doi.org/10.1029/2000GL011893

Lübken, F.-J., Hoffner, J., Viehl, T. P., Becker, E., Latteck, R., Kaifler, B., et al. (2015). Winter/summer transition in the Antarctic mesopaause region. *Journal of Geophysical Research: Atmospheres*, *120*, 12,394–12,409. https://doi.org/10.1002/2015JD023928

Lübken, F.-J., Höffner, J., Viehl, T. P., Kaifler, B., & Morris, R. J. (2011). First measurements of thermal tides in the summer mesopause region at Antarctic latitudes. *Geophysical Research Letters*, *38*, L24806. https://doi.org/10.1029/2011GL050045

Lübken, F.-J., Jarvis, M. J., & Jones, G. O. L. (1999). First in situ temperature measurements at the Antarctic summer mesopause. *Geophysical Research Letters*, *26*(24), 3581–3584. https://doi.org/10.1029/1999GL010719

Meek, C. E., Manson, A. H., Franke, S. J., Singer, W., Hoffmann, P., Clark, R. R., et al. (1996). Global study of northern hemisphere quasi-2-day events in recent summers near 90 km altitude. *Journal of Atmospheric and Terrestrial Physics*, *58*(13), 1401–1411. https://doi.org/10.1016/0021-9169(95)00120-4

Mertens, C. J., Mlynczak, M. G., López-Puertas, M., Wintersteiner, P. P., Picard, R. H., Winick, J. R., et al. (2001). Retrieval of mesospheric and lower thermospheric kinetic temperature from measurements of CO2 15 μm Earth limb emission under non-LTE conditions. *Geophysical Research Letters*, *28*(7), 1391–1394. https://doi.org/10.1029/2000GL012189

Mertens, C. J., Russell, J. M. III, Mlynczak, M. G., She, C.-Y., Schmidlin, F. J., Goldberg, R. A., et al. (2009). Kinetic temperature and carbon dioxide from broadband infrared limb emission measurements taken from the TIMED/SABER instrument. *Advances in Space Research*, *43*(1), 15–27. https://doi.org/10.1016/j.asr.2008.04.017

Mertens, C. J., Schmidlin, F. J., Goldberg, R. A., Remsberg, E. E., Pesnell, W. D., Russell, J. M. III, et al. (2004). SABER observations of mesospheric temperatures and comparisons with falling sphere measurements taken during the 2002 summer MaCWAVE campaign. *Geophysical Research Letters*, *31*, L03105. https://doi.org/10.1029/2003GL018605

Mlynczak, M. G. (1997). Energetics of the mesosphere and lower thermosphere and the SABER experiment. *Advances in Space Research*, *20*(6), 1177–1183. https://doi.org/10.1016/S0273-1177(97)00769-2

Morris, R. J., Höffner, J., Lübken, F.-J., Viehl, T. P., Kaifler, B., & Klekociuk, A. R. (2012). Experimental evidence of a stratospheric circulation influence on mesospheric temperatures and ice-particles during the 2010–2011 austral summer at 69°S. *Journal of Atmospheric and Solar - Terrestrial Physics*, *89*, 54–61. https://doi.org/10.1016/j.jastp/2012.08.007

Muller, H. G., & Nelson, L. (1978). A travelling quasi 2-day wave in the meteor region. *Journal of Atmospheric and Terrestrial Physics*, *40*(6), 761–766. https://doi.org/10.1016/0021-9169(78)90136-8

Panka, P., Kutepov, A. A., Kalogerakis, K. S., Janches, D., Russell, J. M. III, Rezac, L., et al. (2017). Resolving the mesospheric nighttime 4.3 μm emission puzzle: Comparison of the $CO_2(\nu_3)$ and OH (ν) emission models. *Atmospheric Chemistry and Physics*, 17–16. https://doi.org/10.5194/acp-17-9751-2017

Plane, J. M. C. (2003). Atmospheric chemistry of meteoric metals. *Chemical Reviews*, *103*(12), 4963–4984. https://doi.org/10.1021/cr0205309

Plane, J. M. C., Feng, W., & Dawkins, E. C. M. (2015). The mesosphere and metals: Chemistry and changes. *Chemical Reviews*, *115*(10), 4497–4541. https://doi.org/10.1021/cr500501m

Pogoreltsev, A. (1999). Simulation of planetary waves and their influence on the zonally averaged circulation in the middle atmosphere. *Earth, Planet and Space*, *51*(7-8), 773–784. https://doi.org/10.1186/BF03353236

Rapp, M., Lübken, F.-J., Müllemann, A., Thomas, G. E., & Jensen, E. J. (2002). Small scale temperature variations in the vicinity of NLC: Experimental and model results. *Journal of Geophysical Research*, *107*(D19), 4392. https://doi.org/10.1029/2001JD001241

Remsberg, E. E., Marshall, B. T., Garcia-Comas, M., Krueger, D., Lingenfelser, G. S., Martin-Torres, J., et al. (2008). Assessment of the quality of the version 1.07 temperature-versus-pressure profiles of the middle atmosphere from TIMED/SABER. *Journal of Geophysical Research*, *113*, D17101. https://doi.org/10.1029/2008JD010013

Rezac, L., Jian, Y., Yue, J., Russell, J. M. III, Kutepov, A., Garcia, R., et al. (2015). Validation of the global distribution of CO2 volume mixing ratio in the mesosphere and lower thermosphere from SABER. *Journal of Geophysical Research: Atmospheres*, *120*, 12,067–12,081. https://doi.org/10.1002/2015JD023955

Rezac, L., Kutepov, A., Russell, J. M. III, Feofilov, A. G., Yue, J., & Goldberg, R. A. (2015). Simultaneous retrieval of T (p) and CO 2 VMR from two-channel non-LTE limb radiances and application to daytime SABER/TIMED measurements. *Journal of Atmospheric and Solar - Terrestrial Physics*, *130-131*, 23–42. https://doi.org/10.1016/j.jastp.2015.05.004

Rezac, L., Kutepov, A. A., Feofilov, A. G., & Russell, J. M. III (2011). On limb reference calculations and convergence of relaxation type retrieval algorithms. *Applied Optics*, *50*(28), 5499–5502. https://doi.org/10.1384/AO.50.005499

Rodgers, C. D., & Prata, A. (1981). Evidence for a traveling 2-day wave in the middle atmosphere. *Journal of Geophysical Research*, *86*(C10), 9661–9664. https://doi.org/10.1029/JC086iC10p09661

Russell, J. M. III, Mlynczak, M. G., Gordley, L. L., Tansock, J., & Esplin, R. (1999). An overview of the SABER experiment and preliminary calibration results. *Proceedings of SPIE*, *3756*, 277. https://doi.org/10.1117/12.366382

Sharma, R. D. (2015). Technical note: On the possibly missing mechanism of 15 μm emission in the mesosphere-lower thermosphere (MLT). *Atmospheric Chemistry and Physics*, *15*(4), 1661–1667. https://doi.org/10.5194/acp-15-1661-2015

She, C.-Y., Krueger, D. A., & Yuan, T. (2015). Long-term midlatitude mesopause region temperature trend deduced from quarter century (1990–2014) Na lidar observations. *Annales Geophysicae*, *33*(3), 363–369. https://doi.org/10.5194/angeocom-33-363-2015

Stevens, M. H., Lieberman, R. S., Siskind, D. E., McCormack, J. P., Hervig, M. E., & Englert, C. R. (2017). Periodicities of polar mesospheric clouds inferred from a meteorological analysis and forecast system. *Journal of Geophysical Research: Atmospheres*, *122*, 4508–4527. https://doi.org/10.1002/2016JD025349

Tan, B., Chu, X., Liu, H.-L., Yamashita, C., & Russell, J. M. III (2012). Zonal-mean global teleconnection from 15 to 110 km derived from SABER and WACCM. *Journal of Geophysical Research*, *117*, D10106. https://doi.org/10.1029/2011JD016750

Walterscheid, R. L., & Christensen, A. B. (2016). Low-latitude gravity wave variances in the mesosphere and lower thermosphere derived from SABER temperature observation and compared with model simulation of waves generated by deep tropical convection. *Journal of Geophysical Research: Atmospheres*, *121*, 11,900–11,912. https://doi.org/10.1002/2016JD024843

Welch, B. L. (1947). The generalization of 'student's' problem when several different population variances are involved. *Biometrika*, *34*(1-2), 28–35. https://doi.org/10.1093/biomet/34.1-2.28

Welch, B. L. (1951). On the comparison of several mean values: An alternative approach. *Biometrika*, *38*(3/4), 330–336. https://doi.org/10.2307/2332579

Yamashita, C., Liu, H. L., & Chu, X. (2010). Responses of mesosphere and lower thermosphere temperatures to gravity wave forcing during stratospheric sudden warming. *Geophysical Research Letters*, *37*, L09803. https://doi.org/10.1029/2009GL042351

Yuan, T., Pautet, P. D., Zhao, Y., Cai, X., Criddle, N. R., Taylor, M. J., & Pendleton, W. R. (2014). Coordinated investigation of midlatitude upper mesospheric temperature inversion layers and the associated gravity wave forcing by Na lidar and Advanced Mesospheric Temperature Mapper in Logan, Utah. *Journal of Geophysical Research: Atmospheres*, *119*, 3756–3769. https://doi.org/10.1002/2013JD020586

Zhao, J., Chu, X., Chen, C., Lu, X., Fong, W., Yu, Z., et al. (2017). Lidar observations of stratospheric gravity waves from 2011 to 2015 at McMurdo (77.84°S, 166.69°E), Antarctica: 1. Vertical wavelengths, periods, and frequency and vertical wave number spectra. *Journal of Geophysical Research: Atmospheres*, *122*, 5041–5062. https://doi.org/10.1002/2016JD026368