



OH airglow observations with two identical spectrometers: benefits of increased data homogeneity in the identification of variations induced by the 11-year solar cycle, the QBO, and other factors

Carsten Schmidt¹, Lisa Küchelbacher¹, Sabine Wüst¹, and Michael Bittner^{1,2}

¹German Remote Sensing Data Center (DFD), German Aerospace Center (DLR), 82234 Oberpfaffenhofen, Germany

²Institute of Physics, University of Augsburg (UNA), 86159 Augsburg, Germany

Correspondence: Carsten Schmidt (carsten.schmidt@dlr.de)

Received: 28 March 2023 – Discussion started: 22 May 2023

Revised: 24 July 2023 – Accepted: 22 August 2023 – Published: 4 October 2023

Abstract. Hydroxyl (OH) radical airglow observations have been performed at the environmental research station “Schneefernerhaus” (UFS; 47.42° N, 10.98° E) since October 2008, with continuous operation since July 2009. The instrumental setup relies on the parallel operation of two identical instruments, each a GRIPS (GRound-based Infrared P-branch Spectrometer), in order to achieve maximum completeness and homogeneity.

After the first decade of observations the acquired time series are evaluated with respect to the main influences on data quality and comparability to those at other sites. Data quality is essentially limited by gaps impacting the completeness. While technical failures are largely excluded by the setup, gaps caused by adverse meteorological conditions can systematically influence estimates of the annual mean. The overall sampling density is high, with nightly mean temperatures obtained for 3382 of 4018 nights of observation (84%), but the average coverage changes throughout the year. This can bias the annual mean up to 0.8 K if not properly accounted for.

Sensitivity studies performed with the two identical instruments and their retrievals show that the comparability between the observations is influenced by the annual and semi-annual cycle as well as the choice of Einstein-*A* coefficients, which influence the estimate of the annual cycle’s amplitude.

A strong 11-year solar signal of 5.9 ± 0.6 K per 100 sfu is identified in the data. The OH temperatures follow the $F_{10.7\text{cm}}$ value with a time lag of 90 ± 65 d. However, the precise value depends on details of the analysis. The highest correlation ($R^2 = 0.91$) is achieved for yearly mean OH temper-

atures averaged around 4 February and the $F_{10.7\text{cm}}$ solar flux leading ahead with 110 d. A prominent 2-year oscillation is identified between 2011 and 2015. This signal is linked to the quasi-biennial oscillation (QBO), leading to a temperature reduction of approximately 1 K during QBO westward phases in 2011, 2013, and 2015 and a respective 1 K increase in 2012 and 2014 during QBO eastward phases. The amplitude of the semiannual cycle shows a similar behavior with the decade’s minimum amplitudes (~ 2.5 – 3 K) retrieved for 2011, 2013, and 2015 and maximum amplitudes observed in 2012 and 2014 (~ 4 K). The signal appears to disappear after 2016 when the solar flux approaches its next minimum. Although it appears as a rather strict 24-month periodicity between 2011 and 2015, spectral analyses show a more or less continuous oscillation with a period of approximately 21 months over the entire time span, which can be interpreted as the result of a nonlinear interaction of the QBO (28 months) with the annual cycle (12 months).

1 Introduction

At least since the study of Roble and Dickinson (1989) it has been known that the middle and upper atmosphere can provide important insights into long-term changes in atmospheric temperatures and dynamics due to an increase in the CO₂ mixing ratio. But the middle atmosphere (10–100 km), especially the mesosphere–lower thermosphere (MLT) region, poses several observational challenges. The number of high-quality observations spanning a period of more than 1

or 2 decades is limited to a few time series worldwide. At the same time the MLT is believed to be a transition region between different atmospheric regimes. Although temperatures of both the mesosphere and the thermosphere are decreasing, the mesospheric cooling leads to a contraction of the atmosphere accompanied by a shift of the vertical temperature structure. Warmer parts of the thermosphere are supposed to descend to lower heights, creating a region at MLT heights where the decrease in temperature is supposedly compensated for by the descending warmer parts (e.g., Laštovička et al., 2008, 2006, and references therein).

Therefore, long-term studies of MLT temperatures are of special importance for a better understanding of future developments of middle atmosphere dynamics. Only a few measurement techniques grant access to this region well above the stratopause and below the ionosphere. Unlike lidar or radar systems, passive remote sensing of natural airglow emissions originating at heights between 80 and 100 km provides a convenient way of exploring the MLT. Systematic ground-based observations of the rotational–vibrational transitions of the hydroxyl (OH) molecule and the derivation of rotational temperatures from these emissions date back to the 1960s (see, e.g., Semenov, 2000). While Semenov (2000) had to combine data from several sites in his analysis, long time series of a single site are available for a few sites, e.g., at the Kjell Henriksen Observatory (KHO; 78.2° N, 16.0° E) since 1980, at the University of Wuppertal (WUP; 51.3° N, 7.2° E) since 1980, at the Australian Antarctic Station Davis (DAV; 68° S, 78° E) since 1995, at El Leoncito (LEO; 31.80° S, 69.29° W) since 1997, at Maimaga (MAI; 63.04° N, 129.51° E) since 1999, at Zvenigorod (ZVE; 55.69° N, 36.77° E) since 2000, and at several other places (see Holmen et al., 2014; Kalicinsky et al., 2018; French et al., 2020; Reisin and Scheer, 2002; Ammosov et al., 2014; Dalin et al., 2020).

A potential lack of homogeneity and completeness poses a particular threat to long-term OH observations. An instrument failure, for example, will initially cause a gap in the time series, reducing the completeness. If it cannot be repaired the instrument can be replaced with a similar new instrument. Even if both instruments are well-calibrated, the successor will most likely differ in terms of sampling rate, throughput, and other details impacting the homogeneity of the time series. The further the start of observations dates back, the more inhomogeneities and large gaps occur (see, e.g., Kalicinsky et al., 2016; Holmen et al., 2014).

As was pointed out by Laštovička and Jelínek (2019) not only data quality but also analysis methodology and natural variability are sometimes underestimated in trend calculations, which can lead to controversial results. In their review and later updates Beig et al. (2003) address solar variability, more precisely the 11-year Schwabe cycle, as the major natural forcing impacting long-term trend estimates. Among other aspects Laštovička and Jelínek (2019) discuss the multiple regression method, frequently applied in or-

der to decompose a time series into its underlying components such as a bias, cyclic components, long-term trend, and noise. They point out that the more sophisticated and/or complex the analysis method is, the more sensitive it will be to data errors. Ultimately, the issues of instrumentation, data quality, natural variability, and methodology led to the foundation of the Network for the Detection Of Mesospheric Change (NDMC; <https://www.wdc.dlr.de/ndmc/>, last access: 22 September 2023) in 2007 with its mission to foster the exchange of existing know-how and the coordinated development of improved observations, analysis techniques, and modeling.

Shortly thereafter, the prototype of a new instrument was put into operation at the environmental research station “Schneefernerhaus” (UFS) in central Europe (47.42° N, 10.98° E) in 2008 (see Schmidt et al., 2013; Bittner et al., 2010). Based on the previous experience of other research groups in the NDMC, routine observations were planned with two identical instruments (each as a backup for the other) from the very beginning. In this study, the first 11 years (July 2009 until June 2020) covering the decade 2010–2019 and solar cycle 24 are analyzed with respect to data quality and natural variability. It specifically addresses the question of whether the direct comparison of the two instruments’ year-long field observations confirms previous estimates of precision, accuracy, and homogeneity, which previously could only be based on laboratory measurements, theoretical calculations, and empirical “best-practice” efforts.

Figure 1 gives an overview of the generic steps involved in acquiring a long-term time series, including the respective specific steps concerning the OH airglow observations discussed in this study. Once the parameter to be observed has been identified (here: OH rotational temperatures, representing kinetic MLT temperatures), an appropriate instrument has to be chosen or developed (here: a grating spectrograph for the bright OH emissions between 1.5 and 1.6 μm). Then, a certain retrieval needs to be applied to the calibrated observational data (here: interpretation of the $P_1(2)$, $P_1(3)$, and $P_1(4)$ rotational lines of the P-branch of the OH(3-1) vibrational transition). This will result in the desired parameter (OH rotational temperature). In order to get a time series of daily, monthly, or yearly mean values for further interpretation several averaging procedures will usually be applied to the initial data, which may have a temporal resolution of only a few seconds or minutes (10, 15, or 30 s in the case of this study). This study focuses on the procedures shown with solid arrows and boxes in Fig. 1 corresponding to retrieval, averaging, and their consequences for later interpretation and comparison to other observations. It concludes with an analysis of the features observed during the first decade of observations.

The structure of the paper is organized as follows: general aspects of the instruments and temperature retrieval are presented in Sect. 2. The main part in Sect. 3 deals with the remaining differences between the individual instruments and

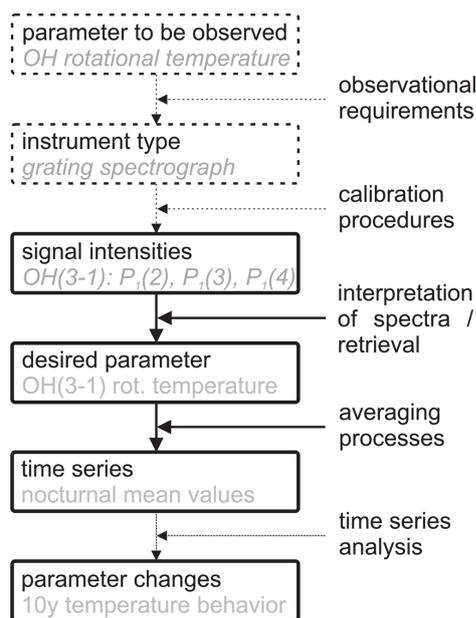


Figure 1. Typical steps involved in designing an operational observation program and respective analyses. The left column shows the general stages (black) and their implementation in this study of OH airglow (gray). Necessary steps to get from one stage to the next one are indicated in the right column. Each step can significantly influence later results. This study focuses on the procedures shown with solid arrows and boxes (that is, retrieval, averaging, and their consequences for later interpretation), and it concludes with an analysis of the features observed during the first decade of observations.

challenges to be met in combining the individual time series. Special emphasis is placed on the non-random appearance of gaps (concerning both nightly and annual mean values). Finally, the time series is used to estimate the solar forcing impact and study a quasi-biennial oscillation (QBO) signal, which seem to be the strongest natural influences impacting the validity of potential future trend estimates. The conclusions to be drawn for future analyses are summarized in Sect. 4.

2 Instrumentation and data reduction methods

2.1 GRIPS

The GROUND-based Infrared P-branch Spectrometer (GRIPS) is designed to separate the P_1 lines from the neighboring P_2 lines of the OH(3-1) rotational vibrational transition for the derivation of the respective rotational temperature from the P_1 lines (that is, a resolving power of $\lambda/\Delta\lambda$ of ~ 500 or a full width at half maximum of 3.1 nm at 1550 nm). All three instruments used in this study are equipped with an ANDOR (Oxford Instruments) spectrograph (Shamrock 163, F/3.6 aperture) and a thermoelectrically cooled InGaAs photodiode array (iDus InGaAs 1.7) with 512 pixels.

More details concerning the setup and data reduction schemes of the instruments are given by Schmidt et al. (2013). A few differences concern the fact that the instruments at UFS are operated with an oblique field of view (FoV) and varying temporal resolution. In their standard setup the instruments point southward at a zenith angle of 45° . Due to their entrance angle of approximately $15.5^\circ \times 15.5^\circ$ this corresponds to an FoV of a few hundred square kilometers at 86–87 km, the height of the emitting layer (see, e.g., Wüst et al., 2020, 2017; von Savigny et al., 2012; Baker and Stair, 1988). The observational filter introduced by spatial averaging due to the FoV size has been analyzed by Wüst et al. (2016) in their study of gravity wave potential energy. The systematic differences introduced into the long-term time series by the varying FoV are discussed in Sect. 3.1 of this study. Figure 2 includes horizontal bars indicating which FoV was used at the respective time (FoV 1: 180° azimuth, 45° zenith angle; FoV 2: $180/61^\circ$; FoV 3: $124/57^\circ$). The standard temporal resolution of all GRIPS instruments is 15 s, but GRIPS 7 outperforms most of the other instruments and is therefore operated at 10 s resolution, while in the early setup GRIPS 5 suffered from a minor misalignment and its resolution was reduced to 30 s until it was replaced with GRIPS 8. Thus, depending on season and temporal resolution the number of spectra obtained each night ranges between 800 and 5100.

Due to their large entrance angle the precise calibration of the instruments is a challenging task, as the entire field of view needs to be homogeneously illuminated. Originally, a quartz halogen lamp (Bentham Instruments Ltd. CL2) and a calibrated transmission diffuser (SphereOptics Ltd. SG 3210) were used. In a joint effort with the Physikalisch-Technische Bundesanstalt (PTB, Germany's national institute of standards) ongoing efforts to improve the calibration chain began in 2015. Meanwhile, a substantial number of results have been obtained, and an adequate presentation thereof is beyond the scope of this study. But it should be noted that the GRIPS instruments outperform the original calibration equipment (halogen lamps) concerning long-term stability. Since a calibration offset between GRIPS 7 and GRIPS 8 of 1.83 K was only precisely identified after years of parallel observation, this offset is kept to emphasize the fact that the instruments are independently calibrated. It is of course corrected for once the data from the individual instruments are combined.

Observations started on 25 October 2008 with a prototype setup of GRIPS 5. Although the instrumental setup was not changed since mid-December 2008, autonomous observations frequently failed until June 2009. Therefore, general data quality is considered sufficiently high for long-term analyses since July 2009. Measurements using GRIPS 7 have complemented these observations since 20 January 2010. GRIPS 5 was replaced with GRIPS 8 on 2 September 2010. While the layout of GRIPS 7 was changed for a couple of sensitivity studies, GRIPS 8 has remained unchanged since

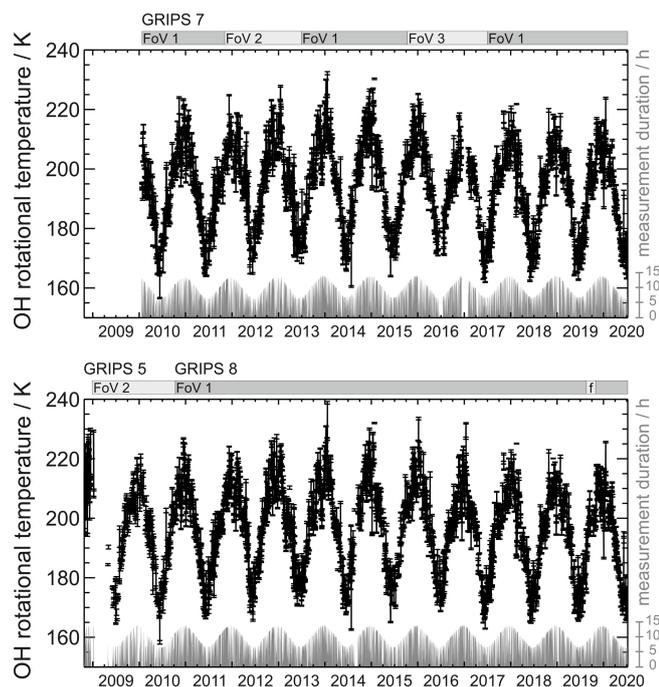


Figure 2. Nightly mean OH temperatures obtained at the UFS “Schneefernerhaus” since October 2008 (black) and respective measurement duration (gray). The upper panel shows data obtained with GRIPS 7, and the lower panel shows data obtained with GRIPS 5 and GRIPS 8. The gray bars above each panel indicate time spans of different fields of view (FoVs). The letter “f” marks a technical failure, which resulted in flawed values of GRIPS 8 during a couple of weeks in 2019; see Table A1 for a comprehensive overview of instrument performance.

the beginning – thereby providing a stable reference. A good overview of the instruments’ performance is provided by their effective observation time displayed at the bottom of the graphs (gray area) in Fig. 2.

2.2 Raw data reduction

The first step in determining the temperature of the OH layer involves applying the physical relationship between the emission intensities $I_{v',J',i \rightarrow v'',J'',i}$ and the OH molecules’ rotational temperature T_{rot} . It was already given by Meinel (1950):

$$I_{v',J',i \rightarrow v'',J'',i} = N_{v'} \cdot A_{v',J',i \rightarrow v'',J'',i} \cdot \frac{2(2J' + 1)}{Q_{v',T_{\text{rot}}}} \cdot \exp\left(\frac{-F_{v',J',i}}{k_{\text{B}} T_{\text{rot}}}\right). \quad (1)$$

Here, v' (v'') and J' (J'') denote the vibrational level and the quantum number of angular momentum of the upper (lower) state of the molecule with i being the doublet branch under investigation. The transition observed with GRIPS is OH(3–1)–P₁ ($v' = 3$, $v'' = 1$, $i = 1$ /electronic state $X^2\Pi_{3/2}$). A ,

$F_{v',J',i}$, and k_{B} denote physical constants: A , the Einstein coefficient of spontaneous emission; F , the term value of the rotational level; and k_{B} , the Boltzmann constant. In principle, it is possible to derive T_{rot} from a single emission line described by Eq. (1), given that a high-quality absolute calibration is available for the observed emission intensities $I_{(v',J',i \rightarrow v'',J'',i)}$. However, the required accuracy and precision of the calibration are hard to achieve, and in addition the population number $N_{v'}$ and the state sum $Q_{v',T_{\text{rot}}}$ are not readily accessible but constant for lines of a single branch.

Therefore, the first three lines P₁(2), P₁(3), and P₁(4) of the $X^2\Pi_{3/2}$ state are observed and a set of equations is formulated and linearized by taking the logarithm. It takes the shape of an inverse problem: $\mathbf{y} = \mathbf{G} \times \mathbf{x}$, with observations (line intensities) summarized in vector \mathbf{y} (dimension: $1 \times n$), unknown properties in parameter vector \mathbf{x} ($m \times 1$), and the physical relationship between the two summarized in matrix \mathbf{G} ($n \times m$); n refers to the number of lines observed (here: $n = 3$) and m to the number of unknown quantities (here: $m = 2$ for T_{rot} and the ratio $N_{v'}/Q_{v',T_{\text{rot}}}$). This inverse problem is solved for T_{rot} by applying the least-squares method, leading to

$$\mathbf{x} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{y}. \quad (2)$$

Although applying the least-squares method is considered straightforward in this case, it implicitly attributes more weight to emission lines which are further apart in the spectrum (here: P₁(2) and P₁(4)). Retrieving correct estimates for the intensities of these two lines from the observed spectra is therefore crucial. Since the P₁ lines partially overlap with the P₂ lines (of the $X^2\Pi_{1/2}$ state), line intensities are estimated by only considering the intensities at the line centers. A correction is still required as P₁(4) is close to R₁(6) of the OH(4–2) transition (1543.22 nm vs. 1543.02 nm, approximately 1 order of magnitude smaller than the spectral resolution of GRIPS). Their overlap is temperature-dependent, and an iterative correction algorithm was provided by Lange (1982) and re-evaluated by Schmidt (2016). The results of their studies (which may be difficult to access) can be summarized as follows: at 150 K the correction amounts to -1 K and is exponentially increasing for increasing temperatures with -3 K at 200 K and -13 K at 250 K. Alternative approaches indicate a precision of ± 0.5 K for this correction, with the overall performance decreasing beyond 230 K and failing to converge beyond 283 K (see also Fig. 3).

2.3 Einstein-A coefficients in the retrieval

According to Eq. (1) the relationship between emission intensities and rotational temperature requires knowledge of several physical constants. But unlike the precisely known Boltzmann constant, k_{B} , several sets of constants have been published for both Einstein-A coefficients of spontaneous emission and rotational term values F . The Einstein-A coefficients of the hydroxyl transitions have previously been

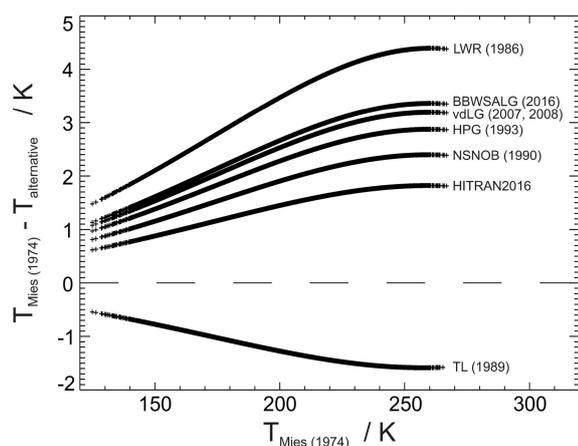


Figure 3. Differences between rotational temperatures if different sets of Einstein-A coefficients are applied during the (otherwise identical) retrieval. The temperatures are obtained from more than 20 000 spectra obtained with GRIPS 8 during cold (summer) and warm (winter) conditions. The temperatures of this study have been calculated with Einstein coefficients published by Mies (1974) unless otherwise noted. BBWSALG – Brooke et al. (2016), HPG – Holtzclaw et al. (1993), LWR – Langhoff et al. (1986), NSNOB – Nelson et al. (1990), TL – Turnbull and Lowe (1989), vdLG – van der Loo and Groenenboom (2007, 2008), HITRAN – High Resolution Transmission molecular spectroscopic database (Gordon et al., 2017; Goldman et al., 1998). The linear relationship is evident from 130 to 230 K, but a deviation from it occurs at higher temperatures. This is due to the correction algorithm discussed in Sect. 2.2.

identified to be of great importance for certain OH vibrational branches (e.g., Noll et al., 2020; Liu et al., 2015). In their analysis of the OH(6-2) transition French et al. (2000) found that temperatures can differ by up to 13 K depending on the set of Einstein-A coefficients.

The influence of the Einstein-A coefficients in the GRIPS retrieval is further investigated in more than 20 000 observed spectra of varying temperature and signal-to-noise ratios of GRIPS 8. Figure 3 shows the temperatures retrieved from these observation data based on eight different sets of Einstein coefficients, including the coldest (warmest) summer (winter) periods observed between 2010 and 2020. The figure displays the differences from temperatures retrieved with the coefficients of Mies (1974) – the standard coefficients still used in the GRIPS retrieval. An important property concerns the fact that the Mies (1974) coefficients lead to higher temperatures than the more recent coefficients of Brooke et al. (2016), van der Loo and Groenenboom (2008), or HITRAN 2016 (High Resolution Transmission molecular spectroscopic database, Gordon et al., 2017; Goldman et al., 1998). Moreover, the differences depend on the absolute temperatures and thus seasons: in winter (at ~ 230 K) the differences are twice as large as in summer (at ~ 160 K). This influences further comparisons, as Mies (1974) coeffi-

cients will lead to systematically higher annual amplitudes than those from Brooke et al. (2016).

However, the temperatures based on the different coefficients can be exactly converted from one set to the other, as these retrievals are not influenced by the signal-to-noise ratio of the observations: each “line” in Fig. 3 consists of more than 20 000 individual data points. Actually, the differences exhibit a linear relationship; the curvature at higher temperatures is due to the decreasing accuracy of the correction algorithm outlined in Sect. 2.2.

Unless otherwise noted, the temperatures from the GRIPS instruments are still based on the Einstein-A coefficients from Mies (1974) and rotational term values taken from Krassovsky et al. (1962). The latter are less often discussed because the influence of different sets of term values is smaller by 1 order of magnitude. Concerning the GRIPS retrieval only the values of Kvifte (1959) may lead to differences of up to ± 0.4 K. Other sets of term values lead to temperatures which are systematically lower but no more than -0.1 K. In general, the respective differences exhibit a smaller temperature dependence and are to some extent influenced by the signal-to-noise ratio; see also Fig. A1 in Appendix A. In contrast to the Einstein-A coefficients employed, the influence of the term values on the retrieval is negligible and not discussed any further here.

3 Results and discussion

3.1 Combining data from different instruments and sites

Occasionally, the data completeness at a measurement site or the sampling rate of a sole instrument is insufficient for an intended analysis of longer-term variability. Thus, time series from different instruments or sites have to be combined. Semenov (2000) was among the first to combine OH temperature time series from several sites (even continents) acquired by different operators with different instruments and retrievals. Despite many corrections applied before comparing these data this can easily compromise the homogeneity of the time series under investigation. Since the pioneering work of Semenov many lessons have been learned and better corrections are usually applied today.

Offermann et al. (2010) combined data from two sites (separated by ~ 500 km) acquired with identical instruments and retrievals. They compared each dataset individually to satellite observations with TIMED-SABER (Thermosphere Ionosphere Mesosphere Energetics Dynamics, Sounding of the Atmosphere using Broadband Emission Radiometry). The satellite-based instrument thus served as a transfer standard. They found that despite some minor differences these datasets agree within their uncertainty estimates. However, their instruments (GRIPS 1 and 2) are predecessors of the instruments used in our study, and the higher precision of

GRIPS 5, 7, and 8 in combination with more than 10 years of parallel observations allows a more detailed investigation of the challenges to be met in combining OH temperature time series. In addition, comparisons between ground- and satellite-based instruments involve certain challenges of their own; e.g., the observed volumes rarely match exactly, leading to the so-called miss-time and miss-distance errors (see, e.g., Reisin and Scheer, 2017; Wüst et al., 2016; Wendt et al., 2013; French and Mulligan, 2010, and references therein).

Since October 2010 the instruments GRIPS 7 and GRIPS 8 have been operated in a parallel setup with identical FoVs. Only twice was the viewing direction of GRIPS 7 changed, when Wüst et al. (2016) studied the impact of the observational filter due to the size of the FoV in gravity wave observations. Therefore, the zenith angle of GRIPS 7 was increased slightly from October 2011 until September 2013 so that it did not overlap with the FoV of GRIPS 8, while keeping the azimuth angle of 180° (southward). Figure 4a shows the respective scatter plot of GRIPS 7 temperatures versus GRIPS 8 temperatures (black) with 415 nights covering almost two seasonal cycles. The inclination of the linear fit (dashed blue) to the data is $0.962 (\pm 0.005)$ and thus deviates slightly from 1.0 (solid red). This inclination is in remarkable agreement with the respective inclination derived from the US Naval Research Laboratory Mass Spectrometer and Incoherent Scatter Radar Extended (NRLMSISE-00) climatology for both latitudes (0.961 ± 0.001), although the absolute temperatures taken from MSISE are lower (dashed green triangles). Figure 4b shows the time period from October 2015 until June 2017 when the setup of GRIPS 7 was changed again, only this time it not only pointed at a different latitude but also at a different longitude further away from the GRIPS 8 FoV. Based on 354 nights the observed temperature dependence (0.981 ± 0.005) does not agree with the one expected from MSISE climatology (1.011 ± 0.001) this time. The remaining parallel alignments summarized in Fig. 4c amount to 1278 nights showing a temperature dependence of 0.991 ± 0.002 (black dots with error bars), which is close to the expected 1.0. The nights shown with green symbols have been excluded from analysis due to more restrictive quality criteria defined below (see Sect. 3.4 and Fig. 6). Both the mean difference of 1.83 K and the 1σ scatter of 1.15 K are significantly smaller when the FoVs agree. The individual parallel alignments (shown in Fig. A2 in Appendix A) indicate that the temperature dependence is reproducible with individual values of 0.995, 0.987, and 0.991.

Figure 5 shows the temperature differences calculated by using the two instruments for the time period, when the GRIPS 7 FoV was changed, so it no longer overlapped with the GRIPS 8 FoV. The mean difference and variability (1σ) increase by approximately 0.7 K (from 1.8 to 2.5 K; note that due to the distribution of gaps the observed difference noted in Table 1 is biased and appears to be higher). Furthermore, the difference shows a seasonal behavior: it is larger from

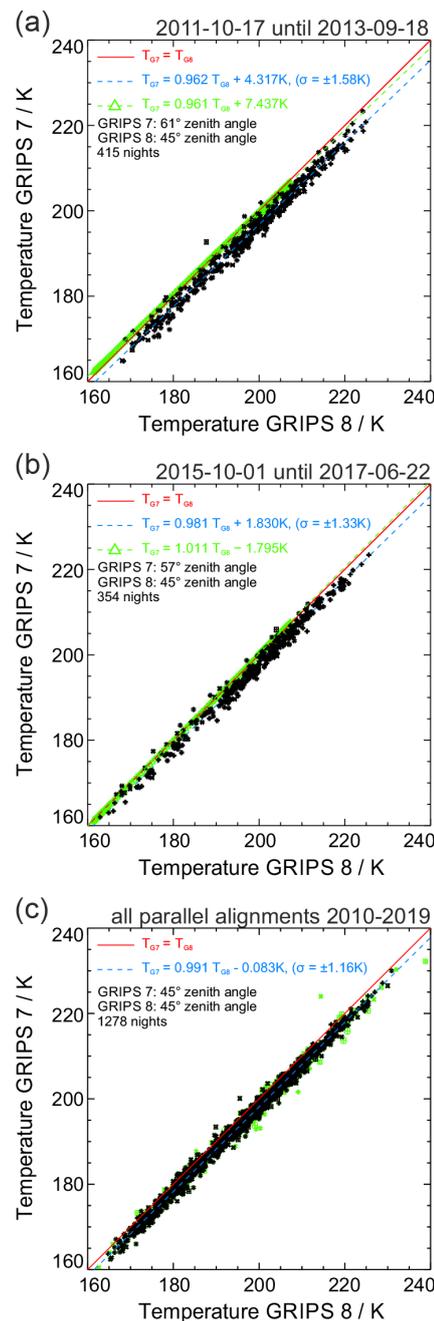


Figure 4. Scatter plots of GRIPS 7 vs. GRIPS 8 derived rotational temperatures for different years (bisecting lines: solid red, fit to the observed data: dashed blue, NRLMSISE-00 climatology values: dashed green triangles). GRIPS 8 stays unchanged with a zenith angle of 45° and azimuth of 180° throughout all years. (a) Both instruments pointing southwards, but GRIPS 7 has a larger zenith angle of 61° ; (b) both zenith and azimuth angles of GRIPS 7 are changed to 57° and 124° , respectively; (c) all parallel alignments between 2010 and 2019 – green data points in panel (c) – are excluded later due to more restrictive criteria than in the original processing scheme (see also Fig. 6). Note: the green triangles in panels (a) and (b) are rather close to each other and appear as one thick green line. Details are discussed in the text.

August to April (up to +1.2 K in October–November) and smaller during summer from May to July (down to −0.2 K in May). The green line shows the differences, which are expected when the NRLMSISE-00 climatological means for the respective positions (at 87 km and local midnight) are adopted. Fitting the annual and semiannual components to the data with a harmonic analysis almost perfectly reproduces the course of the NRLMSISE-00 climatology. Obviously, the seasonal cycle can explain the different slopes observed in the scatter plot in Fig. 4a. Unfortunately, technical failures in 2015 and 2016 caused large gaps (see “measurement duration” in Fig. 2 and Table A1), so no reliable results can be obtained for the second time period when the FoV was changed.

The temperatures obtained from the two instruments are not drifting apart, which is proven by the repeated parallel alignments. Yet, the temperatures begin to show systematic differences once small changes between the setups are introduced (see Table 1). These systematic differences will further increase if data are compared to data from other observing sites. Since these differences are caused by latitudinal differences and their respective seasonal cycles, the temperature difference between two observing sites can hardly be approximated by a simple fixed offset; even a linear relationship as shown in Fig. 4 may improperly represent the differences between two sites. Although Offermann et al. (2010) performed a thorough analysis before combining their data from Wuppertal (51° N, 7° E) and Hohenpeißenberg (48° N, 11° E), the larger variability in their data probably prevented the identification of a seasonal dependence beyond their estimated offset of 0.8 to 1.2 K between the two sites.

3.2 Uncertainty of annual and nightly mean values

Since long-term time series of OH temperatures are usually composed of nightly mean values, the quality of the entire time series depends on both the quality and quantity of the individual nightly mean values. The quantity is easily determined in the case of the UFS time series: once the data from all instruments are combined (with systematic biases shown in Figs. 4 and 5 removed) there are on average 287 valid nightly mean values available per year ($\sim 78.5\%$).

The situation is more difficult concerning the quality of the nightly mean. It is usually expressed via precision ΔT (addressing random scatter) and accuracy, approximated by the bias T_{Bias} to a reference (accounting for systematic errors); in the following, quantities with Δ refer to quantities which increase the random scatter, and those with the subscript B refer to quantities which can have a systematic influence. ΔT is mainly composed of four contributions:

$$\Delta T = \Delta T_{\text{stat}} + \Delta T_{\text{samp}} + \Delta T_{\text{meth}} + \Delta T_{\text{cal}}. \quad (3)$$

Here, ΔT_{stat} is a statistical uncertainty, mainly driven by the number of individual temperature values contributing to the nightly mean value. ΔT_{samp} is the uncertainty due to the ac-

tual sampling over the course of the night (e.g., clouds can significantly influence the value of ΔT_{samp} by obstructing the FoV). Sometimes it is also referred to as “geophysical noise” caused by incomplete averaging of atmospheric waves, especially tides, correlating the individual observations (e.g., Beig et al., 2003). ΔT_{meth} is the uncertainty due to intrinsic properties of the methods applied in the retrieval and is often negligible. An example of ΔT_{meth} coming into play is the exponential relationship between OH line intensities and rotational temperatures: to a small extent higher temperatures will therefore have higher uncertainties (see Eq. 1). ΔT_{cal} represents the precision of the calibration chain from the field calibration of the instrument via secondary standards to some national reference source. These reference sources are radiance or irradiance standards; determining their contribution to ΔT involves sophisticated methods, which is beyond the scope of this study.

Similarly, T_{Bias} depends on several contributions:

$$T_{\text{Bias}} = T_{\text{B, cal}} + T_{\text{B, samp}} + T_{\text{B, meth}} + T_{\text{B, repr}}. \quad (4)$$

Here, $T_{\text{B, cal}}$ is the calibration offset of the instrument, $T_{\text{B, samp}}$ is a potential bias introduced by the actual observational coverage over the course of the night, and $T_{\text{B, meth}}$ is a bias introduced by the methods applied in the retrieval. A well-known contribution to $T_{\text{B, meth}}$ is the mandatory application of Einstein-A coefficients in the retrieval, with different sets yielding different results (see Fig. 3). $T_{\text{B, repr}}$ addresses the question of to what extent OH rotational temperatures are representative of the kinetic temperature of the atmosphere.

The contributions of ΔT_{meth} , ΔT_{cal} and $T_{\text{B, cal}}$, $T_{\text{B, meth}}$, and $T_{\text{B, repr}}$ are assumed to be (more or less) constant over time. The remaining ΔT_{stat} , ΔT_{samp} , and $T_{\text{B, samp}}$ can vary between individual nights. It is especially difficult to estimate ΔT_{samp} and $T_{\text{B, samp}}$. Commonly, the effective observation time during a given night is taken as a criterion to identify nights when contributions of ΔT_{samp} and $T_{\text{B, samp}}$ are negligible. But there is no common agreement concerning how long this effective observation time needs to be. For GRIPS a value of 2 h was empirically derived early in 2009 (Schmidt et al., 2013). Pautet et al. (2014) use a value of at least 4 h, and Reisin and Scheer (2002) use a value of at least 3.6 h – sometimes required to be distributed over an actual observing time of 6 h. Both point out that these are typical values referring to the majority of their observations but that ultimately the averaging needs to meet the requirements of the intended analysis (Dominique Pautet, Esteban Reisin, Jürgen Scheer, personal communication, 2021). Other researchers apply different approaches. Bittner et al. (2002) use a lower limit of at least 1 h of observation time, with a scientist carefully approving each night in addition. A similar statement was brought forward by Lowe: “depending on your site 1 hour of observation time might be sufficient if sampled around local midnight” (Bob Lowe, deceased, personal communication, 2010). The latter statement points to the primary issue of sun-synchronous tides being able to systematically shift the

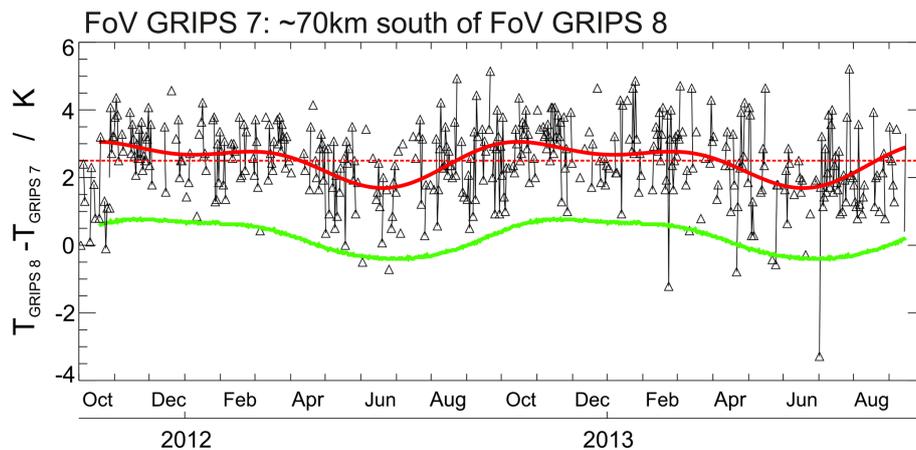


Figure 5. Observed differences between the nocturnal temperatures (black) when the FoV of GRIPS 7 was changed from a zenith angle of 45 to 61° showing the same data as Fig. 4a. The green line shows the differences expected based on NRLMSISE-00 climatological means. The solid red line shows the result of a harmonic analysis, reproducing the annual and semiannual oscillations. The dashed red line shows the (bias-corrected) mean difference.

Table 1. Systematic differences between the temperatures obtained with GRIPS 7 and GRIPS 8. Changing only the effective latitude of the GRIPS 7 FoV (October 2011–September 2013) had the most significant influence on the differences. The repeated parallel alignments show no systematic drift between the instruments. Note: due to repairs affecting the optics of GRIPS 8 this comparison ends in August 2019. Including data extending beyond the time frame of this study, the systematic differences currently remain at a level of 1.77 ± 1.05 K ($0.994 T_{G8} - 0.607$ K).

Observation period	FoV of GRIPS 7 zenith/azimuth	Mean difference $T_{G8} - T_{G7} (\pm 1\sigma)$	$T_{G7} = f(T_{G8})$	Number of nights
2 September 2010 until 16 October 2011	parallel	1.98 ± 1.34 K	$0.995T_{G8} - 1.09$ K	256
17 October 2011 until 18 September 2013	$61/180^\circ$	3.10 ± 1.58 K	$0.962T_{G8} + 4.32$ K	415
18 September 2013 until 30 September 2015	parallel	1.71 ± 1.07 K	$0.987T_{G8} + 0.80$ K	492
1 October 2015 until 22 June 2017	$57/124^\circ$	2.00 ± 1.33 K	$0.981T_{G8} + 1.83$ K	354
22 June 2017 until 2 August 2019	parallel	1.86 ± 1.13 K	$0.991T_{G8} + 0.15$ K	530
All parallel alignments	$45/180^\circ$	1.83 ± 1.16 K	$0.991T_{G8} - 0.08$ K	1278

airglow temperatures depending on the local time of observation. Scheer et al. (1994) were among the first to compensate for tides biasing their means due to incomplete observational coverage by using data from neighboring nights.

By investigating the large dataset comprising 10 years of parallel observations with two identical instruments, it is possible to investigate this sampling issue in greater detail. Figure 6 shows the temperature differences between nightly mean values as a function of the difference in observation time. Each instrument acquired data for at least 2 h. Figure 6a and b show the data during times with identical as well as different FoVs. During the majority of nights, the measurement duration does not differ more than 45 min (see overplotted bar graph). The average observation time of GRIPS 8 is longer because the temporal resolution of GRIPS 8 is 15 s and only 10 s for GRIPS 7. Thus, the signal received by GRIPS 8 stays above the detection threshold longer during adverse conditions.

Both cases show that about 3% to 5% of the nights are outside of the majority of the distribution. The selection criteria can be adjusted by excluding nights with an absolute observation time difference of more than 1.75 h (blue squares), a relative observation time difference of more than 17.5% (yellow diamonds), or an observation time of less than 3 h (red triangles). The remaining black crosses denote the data which have been used in the scatter plots in Fig. 4. Although only 50 nights appear to be outliers in Fig. 6a more than 140 nights are excluded from further analysis by adding this last criterion. Obviously, the remaining differences still range from roughly -1 to $+4.5$ K, which is larger than their calibration offset (ca. 1.83 K) combined with the typical measurement uncertainty ΔT . Note that ΔT is solely based on ΔT_{stat} in the standard processing scheme (v1.0), resulting in less than ± 1 K for each instrument (in 99% of the nights); see Schmidt et al. (2013) for a detailed discussion of this scheme. But the observed scatter in Fig. 6 indicates that the true value of ΔT is often closer to ± 2 K (or ± 2.75 K/ $\sqrt{2}$).

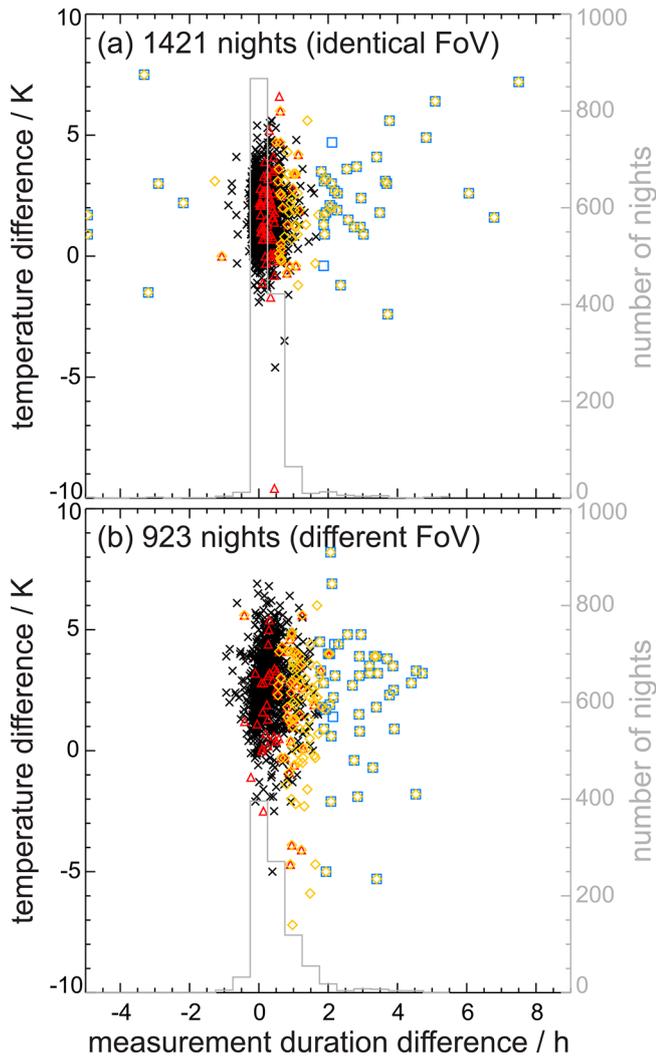


Figure 6. Temperature difference as a function of the instruments' individual measurement duration shown as the observation time of GRIPS 8 minus the observation time of GRIPS 7 for (a) matching FoVs and (b) for separated FoVs. Black crosses denote a minimum observation time of at least 3 h, an absolute difference of less than 1.75 h, and a relative difference of less than 17.5 %; red triangles show nights with observation times of 2 to 3 h; blue squares (yellow diamonds) show nights with observation time differences of more than 1.75 h (17.5 %). During the majority of nights, the observation time agrees within 45 min.

Traditionally, ΔT is almost entirely based on ΔT_{stat} , neglecting contributions from the other sources (given that the selection criteria outlined above are applied). This approach yields reasonable values assuming that the individual temperature values are uncorrelated and exhibit a random scatter, which is larger than the systematic variations caused by atmospheric waves. Then, ΔT can simply be estimated by σ/\sqrt{n} , with the standard deviation, σ , and the number of individual data points, n . However, the instruments' superior temporal resolution makes ΔT_{stat} become fairly small;

GRIPS 7, for instance, acquires more than 700 spectra within 2 h. In addition, statistical approaches usually require the individual data points to be independent. But atmospheric waves change the individual temperatures in a deterministic way.

Figure 7a shows the average course of temperatures during January for the decade 2010 to 2019; Fig. 7b–d show three successive nights observed in January 2018. While the first night shown in Fig. 7b is characterized by almost 8 h of observation time, meeting each of the quality criteria mentioned above, the other two nights actually cover sunset to sunrise. Comparing the first night to the January mean, one should expect its respective mean to be too high by 0.7 % or 1.5 K. But if the other two nights were limited to the observation time of the first night, their retrieved means would actually be too low by 0.7 K (Fig. 7c) or hardly change at all (Fig. 7d). While Fig. 7a shows the influence of sun-synchronous tides, especially the semidiurnal tide, specific nights often show variability due to gravity waves.

Hence, a systematic influence of tides (T_{Bias}) exists, but a respective quantitative estimate thereof also needs to include a proper treatment of the related precision ΔT . The influence of atmospheric waves (tides and gravity waves) on both parameters is therefore based on a statistical approach. For each month the 100 best nights observed in this decade serve as a reference (covering the entire night from sunset to sunrise without gaps). The quality of any new night is now judged by investigating how the retrieved nightly means of the reference nights change once the actual observation time of the observed night is applied to the reference nights. If the gap pattern systematically shifts the mean temperatures of the reference nights towards lower or higher values, then it is causing a systematic sampling bias $T_{\text{B, samp}}$. If the gap pattern only introduces a certain scatter in the nightly means, with some being higher and others being lower than before, then it is only causing an increase in sampling precision ΔT_{samp} . The respective new data processing does not change the initial interpretation of the spectra. Therefore, it is documented as version 1.0A based on the version 1.0 retrieval.

Table 2 shows the respective changes for the three nights shown in Fig. 7. Despite the long time span of almost 8 h, the limitation to the second half of the night causes high estimates for both ΔT_{samp} (± 1.6 K) and $T_{\text{B, samp}}$ (+1.93 K) in the case of the first night. The better the sampling is for a given night, the smaller ΔT_{samp} and $T_{\text{B, samp}}$ become (see the other two nights). The systematic influence of sun-synchronous tides is the main contributor to $T_{\text{B, samp}}$, while ΔT_{samp} is mainly governed by the random influence of gravity waves and non-migrating tides. Note: Fig. 7 displays simple arithmetic means (Table 2: column 2) for easier comparison of these values with the plots. The actual processing involves weighting procedures in both the former data version (column 3) and in the new version (column 4); see Schmidt et al. (2013) for more details.

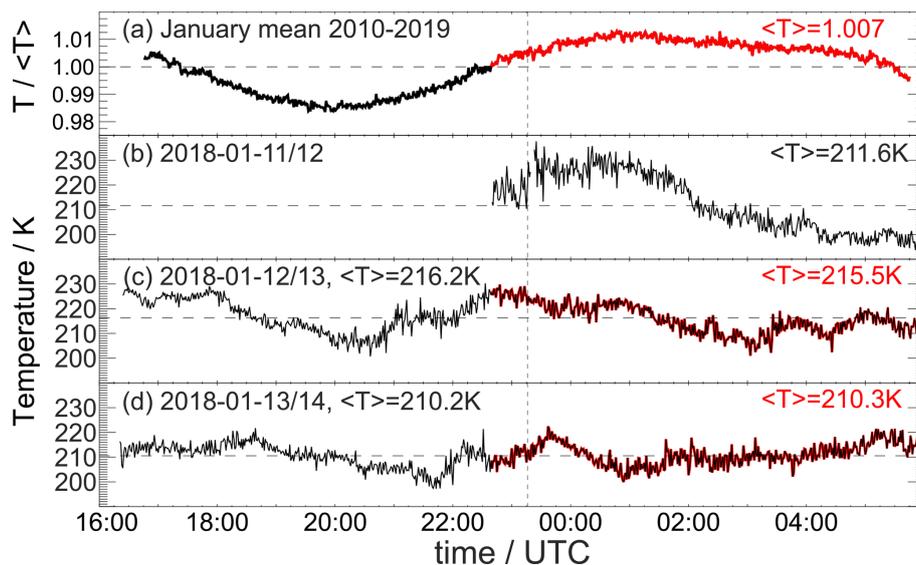


Figure 7. Improving the estimation of the nightly mean. Panel (a) shows the average January night of the decade 2010–2019. The red part highlights the observation time on 11–12 January 2018 shown in panel (b). Therefore, the observed mean value $\langle T \rangle_{\text{obs}}$ of 211.6 K is expected to be too high by 0.7 % ($= 1.5\text{ K}$) for this night. The successive nights shown in panels (c) and (d) indicate that this statistical relationship does not always apply as the respective sampling bias on these nights would actually be -0.7 or $+0.1\text{ K}$. See text for further details.

Table 2. Comparison of the estimates for precision and accuracy of the nights shown in Fig. 7. The actual observing time (length and position within the night) can strongly influence both the sampling bias $T_{\text{B, samp}}$ and the precision ΔT . The last column shows the arithmetic mean of the observation time, indicating whether or not the observations are evenly distributed around local midnight ($\sim 23:17\text{ UTC}$).

	Arithmetic mean (as in Fig. 7)	Classic approach data version 1.0	New approach data version 1.0A	Observing time	$T_{\text{B, samp}}$	Mean time of observation (UTC)
11/12 January 2018	211.6 K	$213.8 \pm 0.1\text{ K}$	$211.8 \pm 1.6\text{ K}$	7.96 h	1.93 K	02:59
12/13 January 2018	216.2 K	$218.6 \pm 0.1\text{ K}$	$218.6 \pm 0.1\text{ K}$	13.68 h	0.01 K	23:49
13/14 January 2018	210.2 K	$212.2 \pm 0.1\text{ K}$	$212.2 \pm 0.1\text{ K}$	13.68 h	0.01 K	23:48

The reassessment of these uncertainties is expanded to the entire dataset. Figure 8 shows nightly means of the years 2014 and 2018. Red triangles and black dots in Fig. 8a and b refer to data retrieved via the former and newer processing, respectively. For the former version, the principal factor in determining the validity of one such nightly mean was the observation time (minimum of 2 h). The improved retrieval now allows for a better assessment of the data quality. In Fig. 8 only nights with an uncertainty of less than 2 % (corresponding to $\Delta T = \pm 4\text{ K}$ at 200 K) are displayed. In this case, the number of nights available for investigation actually increases from 292 nights to 317 nights in 2014 and from 298 to 319 in 2018. If a more restrictive criterion is applied, for instance if only nights with 1 % uncertainty are used, then the number of nights will be slightly smaller than before. On average 307 nights per year (84 %) qualify for further analysis in this improved data version 1.0A if an uncertainty of less than 2 % is required.

While this improvement in error assessment is important for individual nights, the consequences for the annual means are rather small with a change of only 0.1 and 0.2 K. This indicates that nights with positive and negative sampling biases occur rather randomly throughout the time series. It is the expected behavior because incomplete measurement nights are mostly due to bad weather, which on average does not show any systematic occurrence during the night at UFS. Figure 8c and d show the differences between the two retrievals for both years. Although the differences range from $+3$ to -4.5 K , they are indeed more or less randomly distributed and thus have little effect on long-term means. However, there are episodes with specific features that should be kept in mind for certain analyses. During January–February 2014 the sampling bias $T_{\text{B, samp}}$ is often clearly positive for one night and negative for the next one. During April (August) 2014 $T_{\text{B, samp}}$ is systematically high (low). This might be related to tidal activity, but a detailed investigation is beyond the scope of this study. Apparently, several nightly

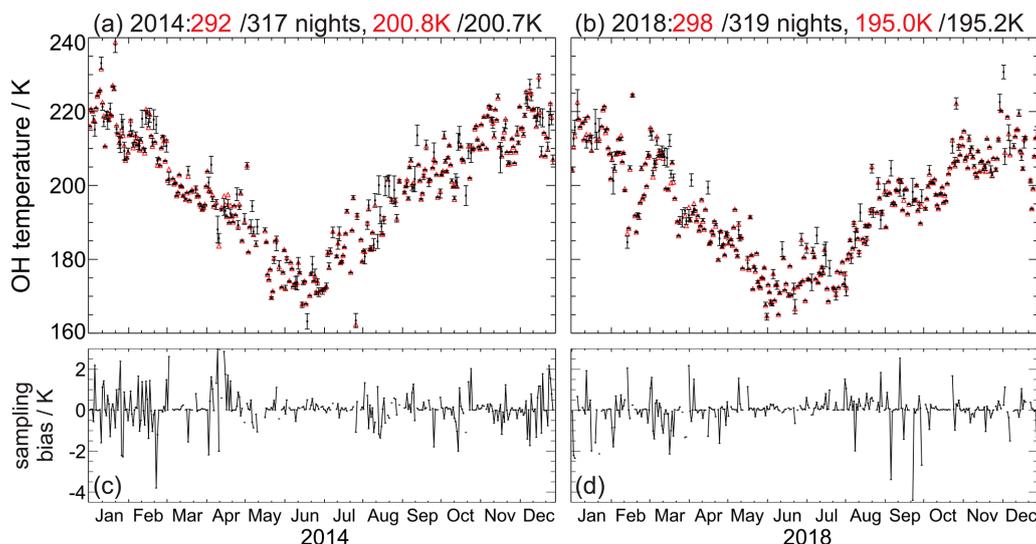


Figure 8. Comparison of the classical (red triangles) retrieval and the improved retrieval (black dots with error bars) for the years 2014 (a, c) and 2018 (b, d). The influence of the retrievals on the annual mean is low because the number of nights with negative or positive sampling biases are almost equal (c, d). However, individual nights can be off by up to 4.5 K (September 2018), and there are episodes with short-term deviations (e.g., April and August 2014). See text for further details.

means lying further away from the majority of the nights are characterized by rather low precision (e.g., low temperatures below 165 K in June–July 2014 or high temperatures above 220 K between October and December 2018). But other “extreme” events such as the temperature variability in February 2018 ranging from 225 to 185 K are clearly of atmospheric origin.

3.3 Annual means and seasonal components

In order to reduce complexity and to remove seasonal features from the data for long-term analyses, annual means are often derived from the time series of nightly means. Similar quality criteria as for nightly mean values apply. Nonuniform sampling during the year in particular can cause a bias of the annual mean. Figure 9 shows the annual means from 2010 to 2019; each panel represents different data reduction schemes, starting with different Einstein-*A* coefficients in panel (a). The differences of the respective annual means are almost constantly 2.6 K (with little variance in the second decimal place). If the remaining gaps are filled before calculating the mean, the annual means will be smaller in 6 out of 10 years (see discussion of gap filling below). This systematic bias is largest in 2015 with 0.8 K. A detailed analysis reveals that gaps are often not uniformly distributed throughout the year, but in May and June there are systematically fewer observations due to prevailing higher cloud coverage in the troposphere. This is the time when temperatures approach the annual minimum (see Fig. 2, “measurement duration”, e.g., in 2011 or 2015). In some years this is compensated for by comparable gaps in winter (when temperatures reach their yearly maximum, e.g., in 2019).

Currently, a dozen sites are equipped with an instrument of this GRIPS type. But UFS is the only site with two instruments operated in parallel to avoid gaps due to technical failures. Thus, it is important to investigate to what extent technical failures change the retrieved annual means and if and how these biases can be corrected. The largest gaps occur in the GRIPS 7 data in 2016 (47 nights) and 2017 (42 nights), changing the mean 2017 by 1.4 K (see Fig. 9b and Table A1). The 2016 mean is less affected, as both low summer values and high winter values are missing. The individual data agree better if all gaps are filled with the respective 10-year mean for each missing night (Fig. 9c). However, in 2017 the GRIPS 7 data still show the largest discrepancy compared to the other time series (0.48 K difference). Since 10-year means (covering high and low temperatures as well as high and low solar activity conditions) may not be available for every observing site or even introduce a new unknown bias, an independent method for filling the gaps is tested: a harmonic analysis (HA; for further information see, for example, Wüst and Bittner, 2006) fits the annual and semiannual components to each year. Annual means are then systematically higher by 0.29 ± 0.08 K as the HA does not reproduce the summer minimum well with these settings. In addition to the HA, the maximum entropy method (MEM) can be used in its capacity as linear prediction filter as described by Bittner et al. (2000). The MEM is an autoregressive method, which regards the current state of a time series as a linear combination of (a finite number of) past states plus a random input. A great benefit of the MEM is that it can be successfully applied to rather short time series. Thus, a missing value can be estimated based on only a few preceding values, while

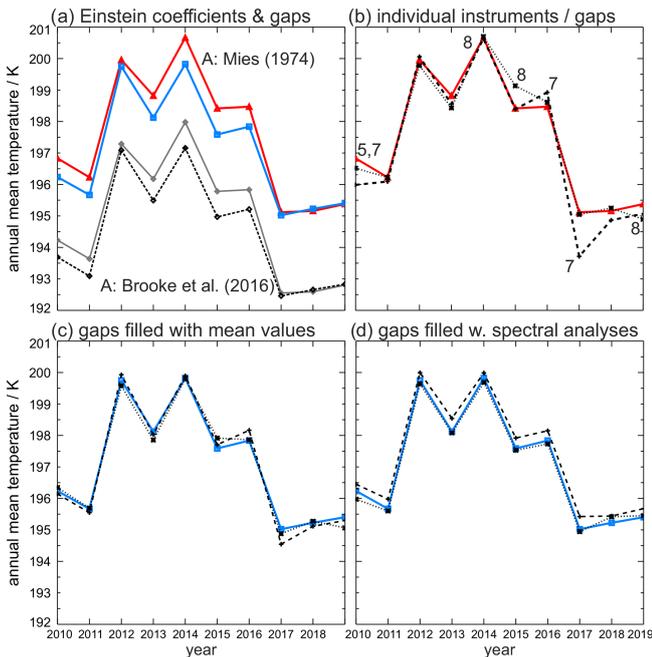


Figure 9. Annual mean OH temperature at the station UFS: solid red curves with triangles and solid blue curves with squares show the same data in all parts. **(a)** Combined data from all instruments; the application of different Einstein-A coefficients in the retrieval leads to a shift of the annual mean (red triangles according to Mies, 1974, vs. gray diamonds based on Brooke et al., 2016; here: 2.6 K). Interpolating the remaining gaps before calculating the mean shifts the annual means up to -0.8 K in some years (solid blue and dashed black). **(b)** Annual means of the individual instruments' data (dashed: GRIPS 7, dotted: GRIPS 5 and 8, red: same as in **a**); the numbers denote the instrument with technical failures during the respective year (see Table A1 for details on large gaps). **(c)** Same as **(b)** but missing nights have been replaced with the decadal mean for the respective nights. Apparently, the influence of the gaps has been eliminated to a large degree (blue: same as in **a**). **(d)** Alternative method of gap filling demonstrated with GRIPS 7 data; gaps have either been filled with a harmonic fit (HA) of the annual and semiannual component each year (dashed) or with the HA and an additional application of the MEM (dotted). The blue line is identical to panels **(a)** and **(c)**.

the properties of the MEM ensure that the first statistical moments (mean and variance) of the time series are not changed. A comprehensive overview of the benefits and caveats of the MEM is given in Bittner et al. (1994). Although the MEM is not designed to fill large gaps, the annual means (dotted line) agree well with the former method (0.06 ± 0.25 K) because the combination of HA and MEM better reproduces the high-frequency variation of the real data (the largest gaps to be filled comprise 14, 12, and 10 nights; see also Table A1). Although a terannual component of the HA is often used for the analysis of OH data (e.g., Ammosov et al., 2014; Bittner et al., 2000), it has not been used here. The terannual component might better represent the existing data but not correctly

estimate the missing data in the case of large gaps. Since the vast majority of observation sites in the NDMC are equipped with only one instrument, the possibility to largely reduce the influence of gaps with these methods is regarded as reassuring.

The seasonal components used in treating gaps are also important parameters in studying atmospheric dynamics. Figure 10 shows amplitudes of the annual and semiannual components (panels a and b) as well as their respective phases (panels c and d). Three different approaches were taken: (1) application of Einstein-A coefficients from Mies (1974) and without interpolation of gaps due to missing nights (solid gray triangles), (2) application of the same coefficients with interpolation of gaps (solid black squares), and (3) application of Einstein-A coefficients from Brooke et al. (2016) with interpolation of gaps (dashed black diamonds). The differences between these three approaches are largest in the assessment of the annual amplitudes. If gaps are not properly treated, the amplitude of the annual component is mostly smaller (up to 0.6 K in 2013) with slightly higher variability. This is due to the fact that in most cases the data gaps that do appear are not at the zero crossings of the annual cycle. The application of different Einstein-A coefficients simply shifts the amplitude by approximately 0.4 K. In general, all three approaches show the same variation, which is characterized by (referring to the solid black line) smaller amplitudes (from 16.8 to 17.2 K) in 2013, 2016, and 2019 and higher amplitudes (from 17.9 to 18.4 K) in 2014, 2015, and 2018. The semiannual amplitude varies between 2.3 and 4.1 K. Einstein coefficients do not influence the results for the semiannual amplitude significantly (less than 0.1 K). But not properly treating the gaps prior to analysis again increases the variability. Between 2011 and 2015 the semiannual component resembles the 2-year periodicity of the annual mean temperatures shown in Fig. 9, with maximum values of ca. 4.1 K in 2012 and 2014 but minimum values of only 2.4 to 2.7 K in 2011, 2013, and 2015. After a hiatus in 2016, the 2-year periodicity appears to continue (further discussed in Sect. 3.4). Figure 10c shows the phases of the annual components. As can be expected from the nature of the annual cycle, the phase is rather constant except in the year 2016, when the phase shifts by 0.15 rad, corresponding to a shift of approximately 9 d. The phases of the semiannual component shown in Fig. 10d exhibit greater variability corresponding to shifts of ± 15 d, but no correlation with other parameters is observed.

3.4 Solar cycle effect and QBO

The correlation between solar activity (widely expressed by the proxy $F_{10.7\text{ cm}}$ index) and OH temperatures has been previously studied by many researchers (see Table 3 for a summary) and high correlation values are evident in the analysis of the UFS datasets, which cover the ~ 11 -year solar cycle 24 almost completely as discussed in this section. Fig-

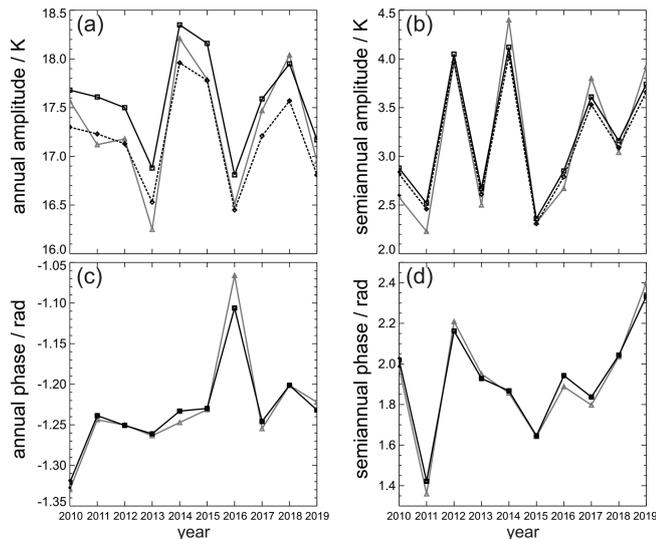


Figure 10. Amplitudes and phases of the annual and semiannual component. **(a)** Amplitude of the annual component based on the retrieval with Einstein-A coefficients taken from Mies (1974) and including gaps (gray triangles) or gaps removed (solid black, squares) and based on retrieval with the Einstein-A coefficients from Brooke et al. (2016) and gaps removed (dashed black, diamonds). The older coefficients lead to a systematically higher amplitude (~ 0.4 K), and interpolating the gaps before the calculation can shift the amplitude by up to 0.6 K. **(b)** The same as **(a)** but for the semiannual component. **(c)** The same as **(a)** but for the phases. **(d)** The same as **(c)** but for the semiannual component. Note: the solid and dashed black curves are almost indistinguishable in panels **(c)** and **(d)**.

Figure 11a shows yearly means of the OH rotational temperatures averaged over different temporal intervals. The black line represents the 365 d moving average, effectively eliminating the annual and shorter harmonic components (input data range: 1 July 2009 until 30 June 2020). The colored lines highlight the yearly means based on either the January–December average (red triangles) or the July–June average (blue squares). There are two striking features: (1) lower temperatures around 195 or 196 K from the beginning of the time series until mid-2011 and again from 2016 to the end of the time series as well as (2) a comparatively strong quasi-biennial component between 2011 and 2015 when temperatures are around 198 to 200 K. It appears that this quasi-biennial component is also present prior to 2011 and beyond 2015, but with a much weaker amplitude (see also Fig. 12). Figure 11b shows the $F_{10.7\text{cm}}$ radio flux (extracted from NASA/GSFC’s OMNI dataset through OMNIWeb: <https://omniweb.gsfc.nasa.gov>, last access: 8 September 2020) in black and a 365 d moving average in blue. High OH temperatures are correlated with a high solar flux, and, remarkably, the quasi-biennial component seems to be weakening with decreasing OH temperatures and decreasing solar activity. Figure 11c shows the QBO wind data at 30 hPa (from Freie Universität (FU) Berlin: <https://www.geo.fu-berlin.de/>

[en/met/ag/strat/produkte/qbo/index.html](https://www.geo.fu-berlin.de/en/met/ag/strat/produkte/qbo/index.html), last access: 6 July 2021).

The correlation between OH temperatures and solar flux is strongly influenced by the quasi-biennial component: in 2011 mean OH temperatures appear colder (-1 K for QBO eastward), while in 2012 they appear warmer ($+1$ K for QBO westward). Also, the solar flux variation shows two distinct maxima in early 2012 and in 2014; both coincide more or less with QBO westward phases. However, the 2-year periodicity of the OH temperatures covers a longer time period and roughly resembles the QBO wind variation but only during increased solar activity from 2011 until 2015 (compare red triangles in Fig. 11a and c).

Studies of the stratospheric wind regime have discussed such a link between the QBO and solar cycle for several decades (e.g., Labitzke, 1987; see the review by Baldwin and Dunkerton, 2005, and references therein). According to Labitzke (2005) the Brewer–Dobson circulation (BDC) is enhanced during solar maximum conditions and westward phase of the QBO, leading to a warmer polar stratosphere and weaker polar vortex. Accordingly, the QBO westward phases observed in 2011, 2013, 2015, and (out of sequence) 2016 correspond to relatively lower MLT temperatures. While the OH temperatures increase by approximately 4 K during solar maximum conditions, the QBO then forces a modulation of ± 1 K onto this temperature increase. The relationship between the QBO and OH temperatures was first studied by Neumann (1990). Nikolashkin et al. (2001) performed a similar study and found a clear anticorrelation between their OH temperatures and the westward phase of the QBO. However, their data may not have been well-suited for this kind of analysis (1201 nights over a period of 9 years, with large gaps, sampled at different observing sites used by Neumann, 1990, and only 152 nights over a period of 6 years used by Nikolashkin et al., 2001). Batista et al. (1994) studied the correlation of the $F_{10.7\text{cm}}$ index with OH(9–4) derived temperatures (and other airglow parameters) also regarding the QBO by fitting harmonics to their observation data extending from 1977 to 1986. Their analysis attributed an amplitude of 0.41 K to the harmonic representing the QBO.

A comprehensive overview of the coupling between the QBO and mid- to high-latitude mesospheric temperatures was given by Espy et al. (2011), who explicitly dealt with the fact that a clear QBO signal cannot be found in many OH temperature time series. They propose that the (tropical) QBO modulates the residual circulation from the winter to summer pole and thus the mesospheric summer temperatures (which cannot be observed via OH airglow at high-latitude sites). Our observed phase relationship with warmer (colder) summertime MLT temperatures during westward (eastward) winds of the QBO matches the explanation proposed by Espy et al. (2011). Although a QBO signal is also present in our summertime data, its amplitude is larger during wintertime: between 2011 and 2015 the year-to-year change in monthly mean temperatures amounts to 1–4 K in July but 6–9 K in

Table 3. Impact of solar activity on OH airglow temperatures. Only studies published within the last decade have been incorporated. Several studies deal with the same dataset but discuss different analysis methods or data periods. Therefore, “location” indicates the underlying time series, while “data period” and “data reduction/analysis methods” give details on how the “solar cycle influence” was estimated. Bold numbers refer to the authors’ recapitulatory conclusions. The abbreviation SSW stands for sudden stratospheric warming, UVES for Ultraviolet and Visual Echelle Spectrograph, and m. r. for multiple regression.

Reference	Location	Solar influence (K per 100 sfu)	Data period	Time lag	Data reduction/analysis methods applied
Holmen et al. (2014)	78° N, 16° E	3.6 ± 4.0	1983–2013	–	winter means (November–February)
		3.7 ± 0.9		–	nightly means
		4.4 ± 2.4		–	monthly means
		5.1 ± 2.7		–	monthly means excl. SSW
		8.8 ± 0.9		65 d	nightly means
		6.4 ± 2.3		65 d	monthly means
		ca. 0		1983–2001	–
10.9	2001–2011	–	winter means (November–February)		
Ammosov et al. (2014)	63° N, 130° E	4.24 ± 1.39	1999–2013	0 months	monthly averages (August–April)
		ca. 7		25 months	monthly averages (August–April)
Perminov et al. (2014)	57° N, 37° E	3.5 ± 0.8	2000–2012	–	nightly means (multiple regression)
Perminov et al. (2018)		4.1 ± 0.5	2000–2016	–	yearly means (both: January–December and July–June, m. r.)
Dalin et al. (2020)	68° S, 78° E	3.35 ± 3.22	2000–2018	0 years	winter (October–March)
French and Klekociuk (2011)		3.77 ± 0.69	1995–2010	–	nightly means (multiple regression)
		3.89 ± 0.69		–	monthly means (m. r.)
		4.79 ± 1.02		–	winter means, mid-April–mid-September (m. r.)
		4.35 ± 0.87		–	yearly means (March–October, m. r.)
		4.63 ± 0.63		160 d	monthly means
4.20 ± 0.66	0	monthly means			
French et al. (2020)	51° N, 7° E	4.30 ± 1.02	1995–2018	–	winter means, mid-April–mid-September (m. r.)
3.5 ± 0.21		–	yearly means		
4.7		1988–2008	–	yearly means (simplified approach)	
Offermann et al. (2010)	51° N, 7° E	1 to 6	1988–2015	–	yearly means (for 11-year subsets of time series)
		3 to 5		–	yearly means
Kalicinsky et al. (2016)	51° N, 7° E	5.0 ± 0.7	1988–2015	–	yearly means (multiple regression no. 1)
		4.1 ± 0.8		–	yearly means (multiple regression no. 2)
Kalicinsky et al. (2018)	51° N, 7° E	3.5 ± 1.5	1988–2017	–	summer means (May–July, m. r.)
		3.5 ± 0.8		–	winter means (November–January, m. r.)
Noll et al. (2017)	25° S, 70° W	4.7 ± 0.4	2000–2015	–	UVES astron. telescope (individual spectra)
This study	47° N, 11° E	5.61	2009–2020	0	yearly means, February–January ($R^2 = 0.61$)
		6.20		110 d	yearly means, August–July ($R^2 = 0.91$)
		5.9 ± 0.6		90 ± 65	concluding evaluation

February and 3–8 K in October. Thus, the QBO-related signal in our data appears to mainly originate from variability at the beginning and at the end of northern hemispheric (NH) winter.

A stable phase relationship between the QBO and annual cycle might be required if the QBO is assumed to cause the observed modulation. Salby and Callaghan (2000) point out that the QBO tends to have shorter (longer) periods during solar maximum (minimum) conditions. This increases

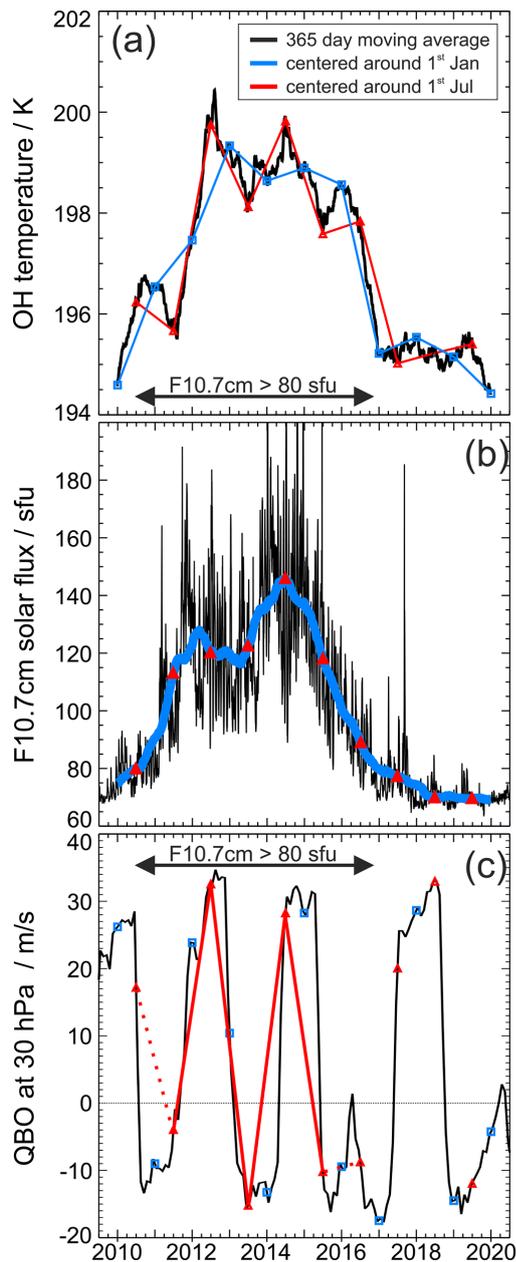


Figure 11. OH rotational temperature, $F_{10.7\text{cm}}$ solar radio flux, and 30 hPa QBO winds (eastward: $< 0\text{ m s}^{-1}$). Panel (a) shows a 365 d moving average of the rotational temperature (black), the January–December yearly means (red triangles; identical to the solid blue curve in Fig. 9), and the July–June yearly means (blue). There is a strong biennial component best seen in the red and black curves. Panel (b) shows daily values of the $F_{10.7\text{cm}}$ radio flux (black, source: NASA/GSFC’s OMNI dataset through OMNIWeb: <https://omniweb.gsfc.nasa.gov>), its 365 d moving average (blue), and January–December annual means (red triangles). Panel (c) shows monthly QBO winds (black, source: <https://www.geo.fu-berlin.de/en/met/ag/strat/produkte/qbo/index.html>). Red triangles and blue squares mark the same dates as in panel (a). Between 2011 and 2015 the phase of the QBO winds more or less matches the biennial cycle (highlighted by the solid red line). See text for further details.

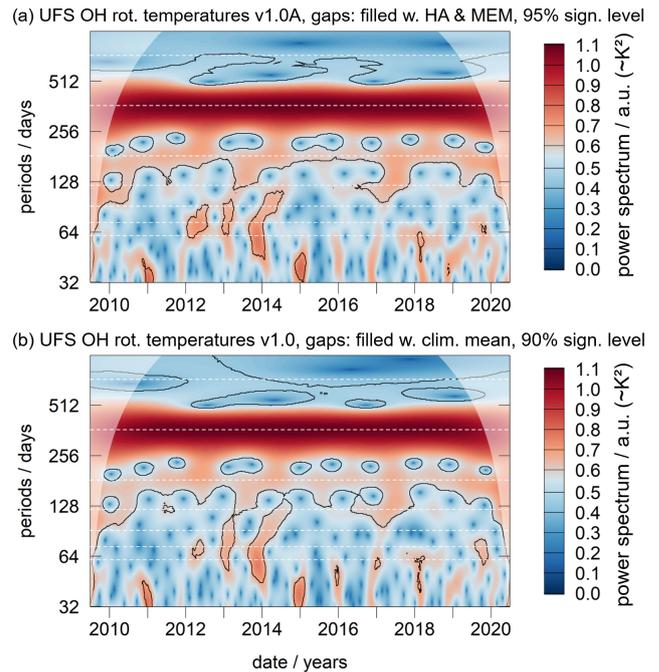


Figure 12. Wavelet power spectrum of the UFS rotational temperatures (shaded areas indicate the cone of influence). (a) Data reduction version 1.0.A; gaps have been filled with a combination HA (366 and 183 d) and MEM, and black lines show the 95 % significance level. (b) Data reduction version 1.0; gaps have been filled with the 10-year mean for the missing date, and black lines show the 90 % significance level (alternative results concerning data version, gap filling, and significance level are shown in Fig. A3 in Appendix A). The dashed white lines highlight periods of 2, 1, 1/2, 1/3, 1/4, 1/5, and 1/6 years. Independent of the analysis, a quasi-biennial component is most significant between 2012 and 2016. However, it may also be present over the entire time at a period of less than 2 years (roughly 650 d or 21.5 months).

the probability of the phases of the two cycles matching for an extended period. Between 2011 and 2015 QBO phases at 30 hPa are indeed fairly stable. NH summers then subsequently fall into eastward and westward phases of the QBO. This relationship breaks in 2015–2016 and 2017–2018, when two subsequent NH summers first fall into the eastward and then into the westward phase of the QBO (see red triangles in Fig. 11c). At the same time the biennial signal in OH temperatures subsides. The QBO pattern observed in OH temperatures from 2011 until 2015–2016 also resembles the pattern of the semiannual amplitude shown in Fig. 10b). Using satellite-based observations Xu et al. (2009) were able to show that the temperature diurnal tide exhibits strong modulations by both the QBO and the semiannual oscillation, especially at heights above 80 km. In addition, the observed period of the temperature tide’s amplitude modulation ranges between approximately 22 and 25 months at these heights. Thus, the multiyear patterns observed in the OH temperatures might all be linked via systematic tidal modulations.

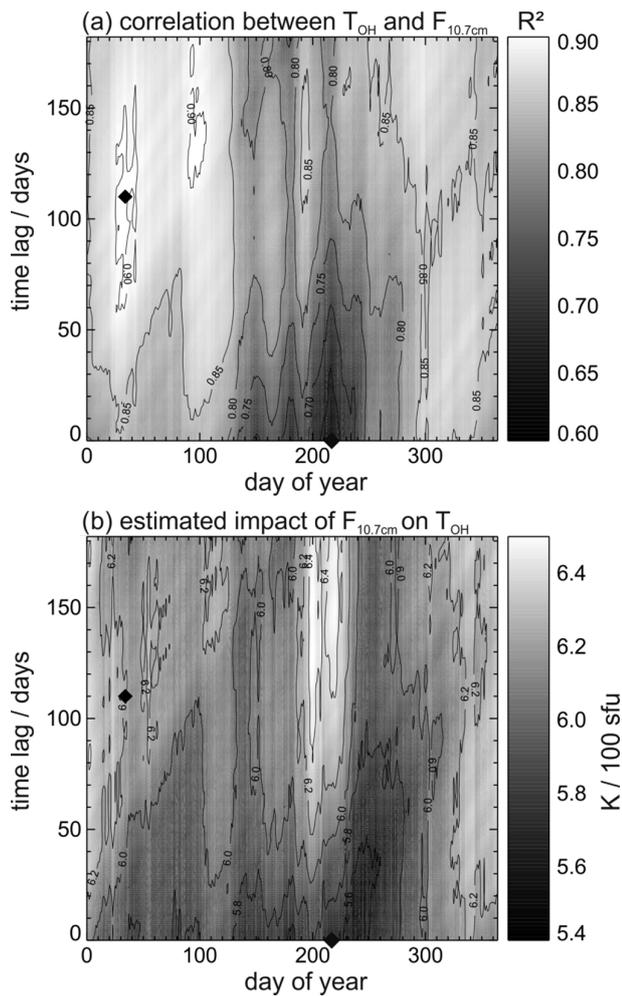


Figure 13. Correlation between the yearly means of T_{OH} and the solar radio flux. Panel (a) shows the correlation (expressed by R^2) as a function of the day of year (doy) and time lag; the highest correlation is achieved when the yearly mean of T_{OH} is centered around day 1 to 120 (January to April) and $F_{10.7cm}$ is assumed to lead T_{OH} by 80 to 160 d. Black diamonds denote days with the highest and the lowest correlation (4 February at 110 d time lag: $R^2 = 0.91$; 6 August, no lag: $R^2 = 0.61$). Panel (b) shows the respective estimates for the impact of the solar flux on T_{OH} . It ranges between 5.4 and 6.45 K per 100 sfu. It amounts to 6.20 K per 100 sfu for the date of the highest correlation and 5.61 K per 100 sfu for the date of the lowest correlation.

Although we observe strong features of the semidiurnal tide during the winter months (see Fig. 7a for January) the discussion of the complex relation between $F_{10.7cm}$, T_{OH} , QBO, semiannual oscillation, and atmospheric tides requires additional data and is beyond the scope of this study.

As gaps have before been interpolated for the correct estimation of the annual means (see Fig. 9), the long-term behavior of the temperatures can now also be analyzed with spectral analysis methods requiring uninterrupted equidistant sampling. The wavelet analysis is such a method, and it

is rather practical for analyzing transient signals. Compared to the short-time Fourier transformation (STFT) it can better produce a representation of the signal in frequency and time because it fits short so-called mother wavelets (wave-like oscillations) to the data at different points in time. Here, we use the Morlet wavelet and the code by Roesch and Schmidbauer (2018) (see “Code and data availability” section). Figure 12 shows two wavelet power spectra of the rotational temperatures for data versions 1.0 and 1.0A, with gaps filled by either the 2010–2019 mean for the respective missing date or alternatively with gaps interpolated by a combination of HA (annual and semiannual amplitude) and MEM in its capacity as a linear prediction filter; shaded areas denote the cone of influence at both ends of the time series. Results for other combinations of data reduction scheme, gap filling, and significance level are largely comparable and are shown in Fig. A3 in Appendix A. Applying these different data reduction and analysis schemes helps discriminate between features which are due to artifacts.

Several features are evident, besides the dominant annual cycle; e.g., the pattern of the semiannual oscillation reproduces the result shown in Fig. 10b well, with the highest amplitudes in 2012 and 2014. A terannual component is present in some years (2013–2014 and 2016–2017) and more or less absent during other years, e.g., in 2018, which shows a strong 60 d oscillation instead (compare also Fig. 8b). The quasi-2-year oscillation can also be identified, but it shows up at periods below 2 years at approximately 650 d or 21.5 months. Note that the time at which this oscillation is above the significance level is subject to the previous data reduction: Fig. 12a shows it above the 95 % significance level from late 2011 until early 2016 (matching the times discussed above). However, it can be considered significant from 2012 until the end of the time series in Fig. 12b with longer periods before 2012. The period of ~ 21.5 months (~ 0.56 cycles per year; cpy) appears to be small. But comparable periods of 0.59 cpy have been reported before by Salby and Callaghan (2000) and Salby et al. (1997) in their investigation of the QBO. They interpret the presence of this oscillation as the result of a nonlinear interaction between the 29-month QBO (0.41 cpy) and the annual cycle (1.0 cpy).

This complex relation between OH rotational temperatures and the QBO regime complicates the study of the relationship between solar flux and OH temperatures. In general, the solar flux appears to lead the OH temperatures by several months because it is increasing in 2011 and already decreasing at the end of 2015. The later increase in OH temperatures in 2012 can be explained by the QBO eastward phase in 2011 causing lower OH temperatures. But this explanation does not apply to the delayed temperature decrease in late 2016: solar flux decreases below 100 solar flux units (sfu) at the end of 2015, and the QBO already returns to eastward winds in mid-2015, while OH temperatures remain high at least until mid-2016.

Clearly, these relationships need further research, but this preliminary examination indicates that multiple regression analyses, which try to simultaneously minimize the influences of the solar cycle, QBO, and other phenomena, might lead to suboptimal results unless they account for a potential intermittency, (nonlinear) interdependency, and lag of some of the forcing mechanisms (see Table 3). Exact results often depend on details of the multiple regression models as was pointed out by Perminov et al. (2018), although Lednyts'kyy et al. (2017) showed how to minimize the degree of freedom by performing a step-by-step multiple linear fit.

A potential lag between solar flux and OH temperatures has been discussed by French and Klekociuk (2011) and Holmen et al. (2014). They found that the correlation between these two parameters will improve if lags of approximately 160 or 65 d are introduced. Ammosov et al. (2014) again briefly discuss the QBO as a potential explanation for their (potential) lag of 25 months. Although rather apparent in their data they do not further emphasize this large lag, lacking a proper mechanism explaining the large delay. All three datasets suffer from the fact that observations have to be suspended for a few months in summer due to the high-latitude observation sites; this clearly complicates certain analyses. However, as diurnal tides appear to be linked to the observed QBO and semiannual oscillations (Xu et al., 2009), either a time-delayed response of these tides or large-scale mixing effects as discussed by Shepherd et al. (2000) might contribute to a time-delayed downward transport and successive heating at OH heights.

The lag between OH temperatures and solar flux is investigated in further detail in Fig. 13. Therefore, the black curve in Fig. 11a is separated into 365 individual time series (i.e., one for each day of the year) composed of 10 data points, each representing a 365 d mean (the red and blue curves in Fig. 11a represent two examples). Thus, the individual data points of each 10-year time series are statistically independent from each other. The same is done for the solar flux. Then the Pearson correlation coefficient, R , is calculated for different time lags between the datasets. Figure 13a shows the value of R^2 as a function of the central date (day of the year around which the OH temperatures were averaged) and time lags of up to half a year (183 d), with the solar flux leading the temperatures. Due to the nature of the data the R^2 values are always high (see Figs. 11 and 14 for comparison). Nevertheless, certain characteristics can be identified: (1) the correlation is higher between November and April ($R^2 > 0.85$), (2) the lowest correlation is obtained for August and small time lags ($R^2 < 0.70$), and (3) the highest correlation ($R^2 > 0.9$) is observed for time lags greater than 60 d in February or April. The highest (lowest) R^2 values of 0.91 (0.61) are obtained if the OH temperatures are averaged around 4 February (6 August) with a time lag of 110 (0) d. Figure 13b shows the respective estimates for the impact of the solar flux on the OH temperatures. These vary between 5.4 and 6.5 K per 100 sfu. Apparently, higher corre-

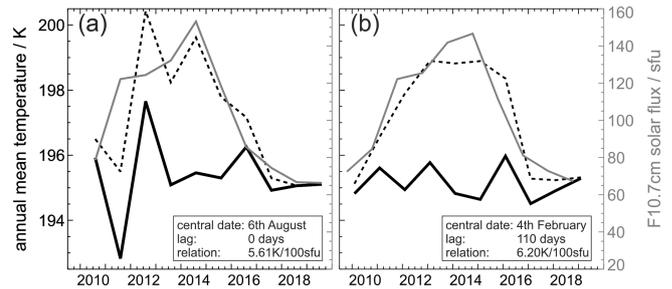


Figure 14. Removing the influence of the solar radiation from the time series of OH temperatures. Panel (a) shows the yearly means of T_{OH} (dashed black) and $F_{10.7\text{cm}}$ (solid gray) centered around 6 August (no lag). The solid black line represents T_{OH} referenced to 70 sfu (assuming 5.61 K per 100 sfu); the 10-year change from 2010 to 2019 then amounts to -0.12 ± 1.42 per 100 K per decade. Panel (b) shows the same as panel (a), but yearly means of T_{OH} (dashed black) and $F_{10.7\text{cm}}$ (solid gray) were centered around 4 February and 17 October (110 d lag), respectively; now the 10-year change amounts to -0.14 ± 0.59 K per decade (and 6.20 K per 100 sfu).

lation implies a higher influence of the solar flux. The highest values of 6.2–6.45 K per 100 sfu are obtained for large time lags (> 70 –170 d) in late July until the end of August, when the correlation reaches only average values between 0.75 and 0.85. It amounts to 6.20 K per 100 sfu on 4 February and 5.61 K per 100 sfu on 6 August. While calculating yearly means eliminates variations with periods up to and below the annual cycle, shifting these averages by half a year minimizes the influence of a potential biennial component. The phase relationship of such a biennial component with respect to the solar flux then determines the value of the observed lag.

For comparison, the time lag and relation between T_{OH} and $F_{10.7\text{cm}}$ are also obtained by calculating the cross-correlation between the 365 d sliding averages of the nightly mean values shown in Fig. 11a and b. The increased number of data points (3652) helps in judging the reliability of the previous results obtained from only 10 well-defined yearly means, although this approach is mathematically not well-defined (since the individual data points are not independent). Here, the results are $5.92 (\pm 0.05)$ up to $6.06 (\pm 0.04)$ K per 100 sfu for time lags between 24 and 85 d with the correlation coefficients showing little variation from 0.780 to 0.782.

In summary, this shows a solar cycle influence of approximately 5.9 ± 0.6 K per 100 sfu at a time lag of 90 ± 65 d. The lag is largely in agreement with the findings of French and Klekociuk (2011) and Holmen et al. (2014). A large uncertainty was adopted for the lag, as it does not appear to be constant over time (see Fig. 13a). If the solar forcing is assumed to be indirectly having an effect on the OH temperatures via changes in middle-atmospheric wind systems and associated waves, this appears to be more than reasonable since these are well-known for their seasonality. This interpretation

is further endorsed by the fact that correlation coefficients are clearly higher from November until April (winter–spring: day 300 to 120 in Fig. 13a). It should be noted that minimum R^2 values are found in August and the minimum forcing is found in September (dark areas in Fig. 13a and b). This is in remarkable agreement with French and Klekociuk (2011), although data and analysis methods differ substantially (see their Fig. 6). Clear evidence for different solar forcing magnitudes in summer and winter was also brought forward by Pertsev and Perminov (2008) and again confirmed by Dalin et al. (2020), with the response in winter being twice as large as in summer. But Kalicinsky et al. (2018), with their observing site being closer to UFS, do not see any difference between winter and summer. A possible explanation lies in the time period covered by the time series. If the forcing is not constant over time, the results may vary because authors are looking at different years. Offermann et al. (2010) argue that results depend on the length of the analysis window and windows should be substantially larger than the studied phenomenon. Holmen et al. (2014), on the other hand, argue that although their analysis window meets this criterion, their result is close to 0 K per 100 sfu for 1983–2001 and +10.9 K per 100 sfu thereafter. Accordingly, they conclude with a reasonably large confidence interval of 3.6 ± 4.0 K per 100 sfu. Table 3 gives an overview of the magnitude of solar forcing on OH temperature time series obtained by various authors during the last decade (dec), including several alternative approaches concerning lags, data averaging, seasonal dependence, and other parameters discussed by the authors.

Although solar cycle 24 is just barely captured by the measurements covered by the UFS dataset and as several circumstances influence the results, there can be no doubt that temperatures roughly increased by 4 K (from 195 to 199 K), in line with a solar flux increase by approximately 70 sfu (see Fig. 11). Regarding the results obtained by other authors, it is questionable if a respective increase of 8 K will be observed for a stronger solar flux increase of 140 sfu, which is frequently found for previous solar cycles. Only future studies can show whether this is (1) due to other long-period oscillations superimposed onto the temperatures as proposed by Höppner and Bittner (2007) and elaborated on by Kalicinsky et al. (2018), (2) due to the solar forcing varying in time (e.g., Holmen et al., 2014), or (3) due to nonlinearity effects limiting the response in the case of stronger forcing.

Ultimately, an appropriate estimate for the solar forcing is required for a precise determination of long-term changes. Figure 14 shows a comparison of two approaches based on the best and worst correlation obtained from Fig. 13a. Figure 14a shows yearly mean temperatures (dashed black) as a result of averaging from 7 to 6 February (centered around 6 August) without any lag to $F_{10.7\text{ cm}}$ (solid gray). Figure 14b shows the same data, but temperature averages are calculated from 6 to 5 August (centered around 4 February) with $F_{10.7\text{ cm}}$ leading the temperatures by 110 d (averaged from 17 to 16 April). The bold black lines result from applying

the respective solar cycle forcing determined from Fig. 13b. Apparently, the resulting overall temperature change in the decade 2010–2019 is comparatively weak in both cases, but the latter approach significantly improves the confidence from -0.12 ± 1.42 to -0.14 ± 0.59 K per decade. This is mainly due to the fact that the temperature deviations in the years 2011 and 2012 are minimized by the latter approach. Since the averaging intervals are shifted by half a year, the latter approach indirectly accounts for the QBO forcing. But this QBO signal is weaker during the last years with reduced solar flux (2017–2019). Therefore, any approach directly addressing the QBO influence (e.g., as part of a multiple regression) should be done with great caution rather than assuming that the QBO influence on OH rotational temperatures was constant over time.

4 Summary and conclusions

A total of 1 decade of airglow observations at the NDMC site UFS were examined with respect to data quality as well as certain aspects of the retrieval and their influence on later results. Advantage was taken of the fact that UFS is equipped with two instruments, which were operated in an identical setup most of the time and with slightly different setups during some years.

The main findings concerning technical aspects, such as data quality and the impact of certain aspects of the retrieval, on the results are the following.

1. The two-instrument setup reliably prevents the occurrence of larger gaps in the dataset due to technical failures of one instrument, and nightly mean values can be obtained for 3382 of 4018 nights. However, the remaining gaps are not randomly distributed in the dataset. At UFS there are systematically fewer observations in May and June due to an increased number of convective clouds. At the same time MLT temperatures reach their annual minimum values. Consequently, estimates of annual or summer mean temperatures must account for this fact. The maximum respective sampling bias for the annual mean is estimated to be +0.8 K in 2015 (see Fig. 9a).
2. Two methods have been applied in filling the remaining gaps: (a) the 10-year average temperature for each missing date and (b) a combination of HA (with 365 and 182.5 d period) and MEM to account for the short-term variability. These perform equally well in correcting the annual mean (compare Fig. 9c and d). In addition, these methods also perform well in correcting the time series of the individual instruments, which exhibit larger gaps due to technical failures (see Fig. 9b and c).
3. An improved method for deriving the nightly mean (and its uncertainty) from a given set of observations was developed in order to better estimate precision, accuracy,

and the probable sampling bias for each night individually. This method is mainly meant to better estimate the influence of sun-synchronous atmospheric tides and avoids empirical constraints (such as a minimum observation time required for a nightly mean to be considered representative). This nightly sampling bias is found to be up to 4 K. However, nights with positive and negative sampling biases are equally distributed throughout the year and there is little influence on the annual mean (± 0.2 K). Shorter averages like monthly means might, however, be influenced by this sampling bias (see Figs. 7, 8).

4. The choice of Einstein-A coefficients is well-known to influence the temperature retrieval. The analysis of all sets of coefficients still in use, from Mies (1974) to Brooke et al. (2016), reveals that the temperature differences depend on the absolute temperature (see Fig. 3). This in turn means that not only estimates of annual means will differ between different retrievals but estimates of the annual cycle's amplitudes will also differ significantly. In the case of the two mentioned sets the annual amplitudes differ by 0.4 K (see Fig. 10).

The improvement of data quality by the parallel operation of two instruments is only a means to an end: increasing the confidence of scientific analyses concerning atmospheric dynamics and especially long-term changes in OH temperatures. The main findings in this context are the following.

1. The amplitude of the annual oscillation is 4 to 6 times larger than the semiannual amplitude (17 to 18.5 K compared 2.5 to 4 K). If the FoVs of both instruments are only slightly separated (~ 100 km), both oscillations will become apparent in the differences of the respective time series (see Fig. 5). This implies that both oscillations experience a strong gradient at UFS (47.42° N, 10.98° E). This needs to be considered when comparing data between sites, which are separated by larger distances, or when data are compared to satellite observations with respective miss distances.
2. Quantitative estimates of the solar cycle response critically depend on the chosen averaging intervals and on the time lags between solar flux and OH temperatures. The highest correlation ($R^2 = 0.91$) is achieved if annual means of OH temperatures are centered around 4 February and solar flux is assumed to lead OH temperatures by 110 d, resulting in a response of 6.20 K per 100 sfu. The lowest correlation ($R^2 = 0.61$, 5.61 K per 100 sfu) is achieved if annual means of OH temperatures are centered around 6 August without any time lag. The cross-correlation analysis of 365 d running means leads to a comparable result of 5.9 to 6.1 K per 100 sfu for time lags between 24 and 85 d ($R^2 = 0.78$). In summary, a solar forcing term of 5.9 ± 0.6 K per 100 sfu covers the observations during solar cycle 24 well.

However, with respect to long-term changes, assuming a response of 6.20 K per 100 sfu and 110 d time lag significantly increases the confidence of the 10-year linear change (2010–2019) from -0.12 ± 1.42 to -0.14 ± 0.59 K per decade (see Fig. 14).

3. Between 2011 and 2016 temperatures show a distinctive quasi-2-year periodicity. During QBO westward phases in 2011, 2013, and 2015 the temperature decreases by approximately -1 K, and during QBO eastward phases it increases by $+1$ K in 2012 and 2014. This appears to be linked to a time when the QBO showed periods closer to 24 months (see Fig. 11). This clear relationship disappears after 2015, when the period of the QBO increases again and solar flux is reduced. But wavelet spectral analyses show a quasi-biennial signal being constantly present, albeit at a lower period of only approximately 21.5 months (~ 0.56 cpy), becoming more pronounced between 2011 and 2016.

This hints at a complex interplay between the solar cycle, QBO, and mid-latitudinal dynamics at MLT heights, with the observed modulation of ~ 0.56 cpy probably being the result of a coupling between the annual cycle (1 cpy) and the climatological average QBO period (~ 0.43 cpy) as proposed by Salby and Callaghan (2000) and Salby et al. (1997). On the other hand, at MLT heights the QBO might only act as a kind of switching mechanism either amplifying or reducing the impact of other forcings which are strictly bound to the annual seasonal cycle.

4. The semiannual cycle shows a similar behavior with the highest amplitudes of 4 K observed in 2012 and 2014 and the lowest amplitudes around 2.5 K observed in 2011, 2013, and 2015. Unlike the biennial signal, the correlation between the QBO phase and amplitude of the semiannual oscillation appears to continue after 2015, with (slightly reduced) amplitudes of 3 K in 2016 and 2018 and slightly increased amplitudes in 2017 and 2019.
5. The year 2016 stands out in several respects: it marks the transition from high to reduced solar activity, and for the first time since 1953 the QBO experienced a disruption of its usual behavior in 2015–2016 (e.g., Newman et al., 2016). It is also remarkable in the UFS OH temperature time series, best seen in the phase of the annual cycle (Fig. 10c), but it also marks the end of the previously pronounced 2-year periodicity mentioned above (black and red curve in Fig. 11a).

The main focus of this study is on data quality aspects, taking advantage of the fact that UFS is equipped with two identical instruments, avoiding large data gaps. The subsequent analysis of the dataset itself deliberately relied on straightforward

methods, revealing strong indications that a QBO-related signal is maximized during enhanced solar activity but less pronounced during the current solar minimum (since 2016). This interdependency of MLT dynamics with the QBO, solar forcing, seasonal cycle, long-term trend, and other phenomena (e.g., La Niña, El Niño) indeed needs more research.

So far, the conservative analyses already performed indicate that at least some forcing mechanisms may have a time-dependent component or exhibit some intermittency not well-represented in common multiple regression approaches. The next steps will therefore include the integration of other observing sites equipped with similar instruments and the application of more sophisticated analysis methods including the variability of airglow intensities.

Appendix A

Table A1. The upper part shows the number of nights with successful data acquisition (including nights with cloud coverage); comments refer to the number marked by an asterisk (*). The lower part lists the remaining large gaps of the final data product, which extend for at least 7 successive days, and the respective cause.

Year	Number of nights		Comment
	GRIPS 5/GRIPS 8	GRIPS 7	
2008	56	–	25 October 2008: first light GRIPS 5
2009	209	–	Many computer-related issues (RAM, USB, etc.)
2010	236/119	340	20 January 2010: start observations with GRIPS 7 1 October 2010: GRIPS 8 replacing GRIPS 5
2011	360	363	–
2012	366	366	–
2013	360	351*	11 September 2013, 9 nights: shutter closed
2014	333*	363	23 August 2014, 25 nights: filter wheel failure
2015	341*	357	2015, 24 nights (intermittent gaps): computer-related issues
2016	363	319*	23 June 2016, 26 nights: computer failure 10 December 2016, 21 nights: filter wheel failure
2017	359	323*	1 January 2017, 29 nights: filter wheel failure contd. + intermittent computer failures
2018	362	359	–
2019	297*	353	11 August 2019, 68 nights: filter wheel failure
2020	364	364	–
Remaining large gaps in final data product			
Year	Cause		Number of missing nights
2009	cloud coverage		8 October 2009: 10 nights
2009	computer related		20 December 2009: 14 nights
2013	UFS power failure		13 August 2013: 7 nights (power failure: 5 nights)
2019	cloud coverage		2 January 2019: 5 and 7 nights (12 of 13 nights missing with one observation night in between)

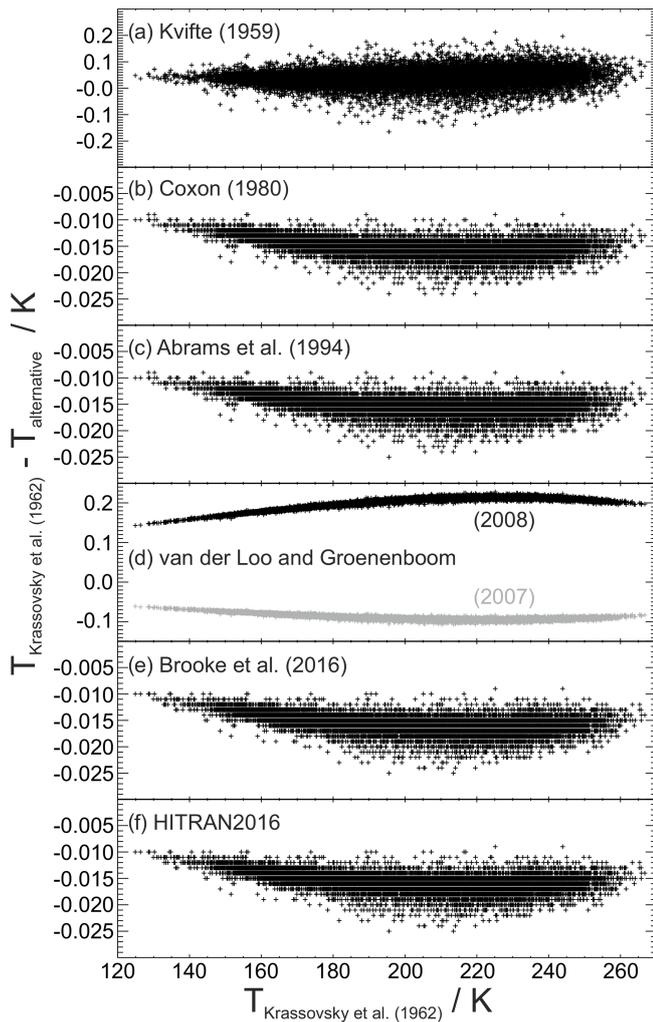


Figure A1. Differences between rotational temperatures if different sets of term values are applied during the (otherwise identical) retrieval. The temperatures are obtained from the same spectra as those shown in Fig. 3. The influence of the term values is 1 order of magnitude smaller than the influence of the Einstein-A coefficients. Note: horizontal white lines in some panels are caused by the limited resolution of 0.001 K.

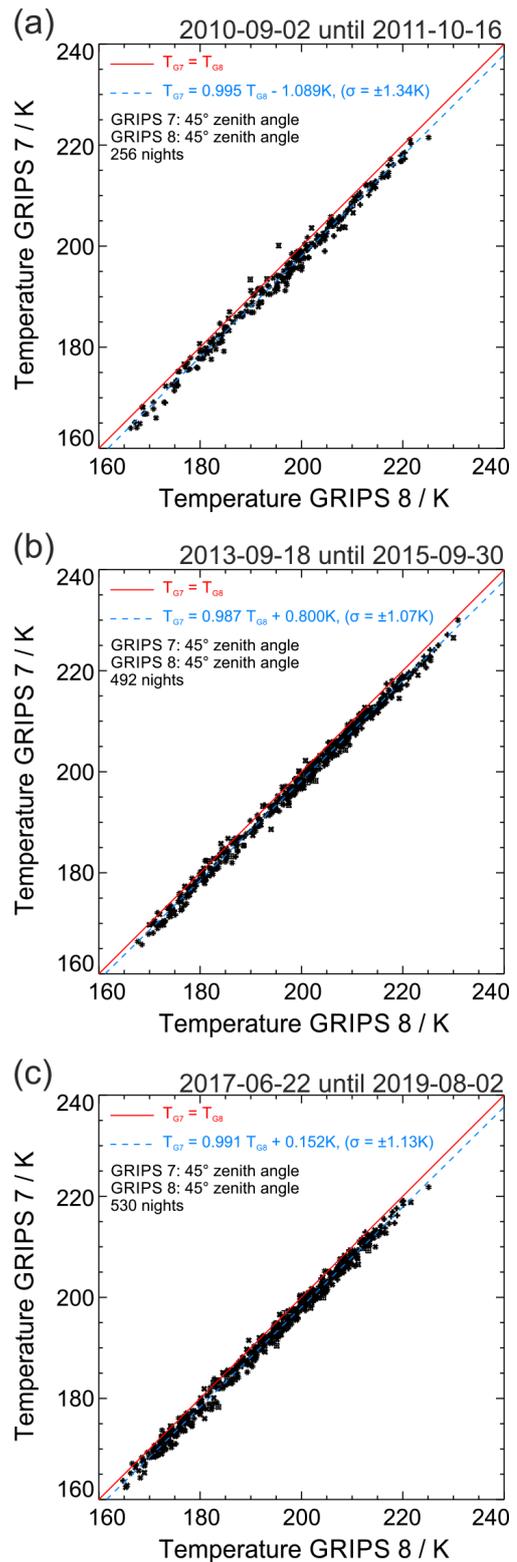


Figure A2. Scatter plots of GRIPS 7 vs. GRIPS 8 temperatures for the individual time periods with matching fields of view of both instruments (bisecting lines: solid red, fit to the data: dashed blue). The mean temperature dependence of $T_{G7} = 0.991 \text{ K} \times T_{G8}$ (see Fig. 4) is reproduced well with individual values of 0.995, 0.987, and 0.991.

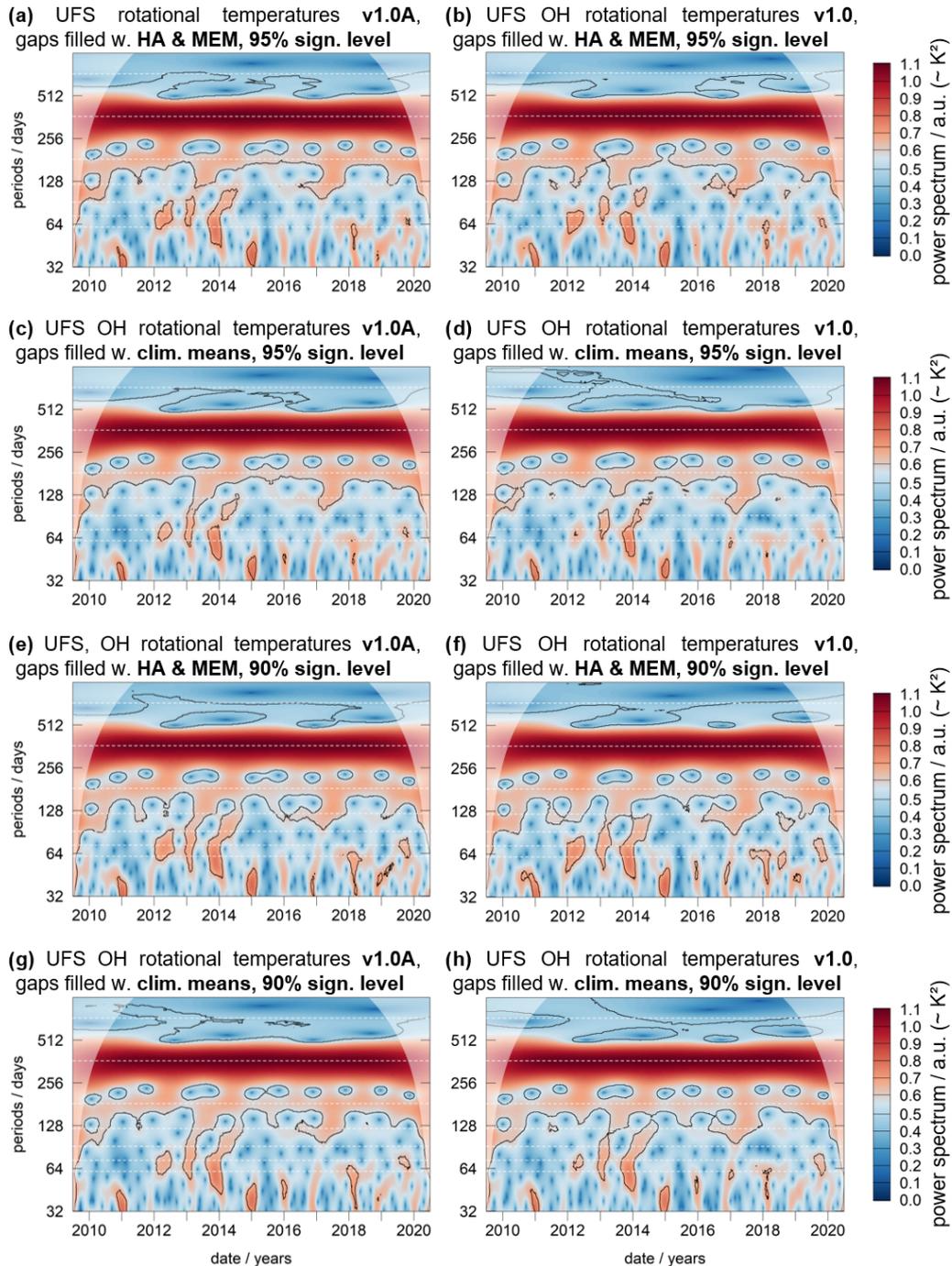


Figure A3. Results of the wavelet spectral analysis for eight possible combinations of time series data version (left column: v1.0A, right column: v1.0), method of gap filling (spectral methods HA and MEM in panels a, b, e, and f or insertion of respective 10-year mean value in panels c, d, g, and h), and significance level of the spectral power (upper four panels: 95 %, lower four panels: 90 %, indicated by the black lines). Shaded areas indicate the cone of influence; panels (a) and (h) are identical to Fig. 12. Most features appear in all of these analyses. A precise identification of the oscillation with a period close to 2 years is difficult but clearly below 24 months (top dashed horizontal white line).

Code and data availability. The OH temperature time series produced and analyzed in this study are available according to the standard conditions outlined by NDMC. Both processing versions v1.0 and v1.0A have been archived at the World Data Center for Remote Sensing of the Atmosphere (WDC-RSAT). They are traceable under the following digital object identifiers (DOIs): <https://doi.org/10.26042/WDCRSAT.XZB5TZQG> (UFS_v10; Schmidt et al., 2023a) and <https://doi.org/10.26042/WDCRSAT.Y0AOE0PZ> (UFS_v10A; Schmidt et al., 2023b).

The wavelet analyses were performed with the R package “WaveletComp”, available at <https://CRAN.R-project.org/package=WaveletComp> (Roesch and Schmidbauer, 2018).

Author contributions. Both the paper and formal analyses contain valuable input from all authors. CS performed most of the analysis; he prepared the original paper including visualization and took care of validation and data curation for more than a decade. LK contributed parts of the paper and performed most of the spectral analyses including visualization of the results. SW contributed parts of the analysis; in particular, she investigated the consequences of changing the FoVs of the instruments. She was also deeply involved in project administration and funding acquisition. MB initiated the observations at UFS including the concept of parallel observations with two instruments. Also, he stimulated the investigation of certain parameters creating an influence during the processing (e.g., Einstein-*A* coefficients). In this role he was responsible for conceptualization, funding acquisition, project administration, and supervision of the observations.

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

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

Acknowledgements. The authors wish to thank Stefan Noll and Oleg Goussev for their valuable input concerning OH Einstein coefficients and Oleg Goussev for assistance with long-term archiving and referencing the data at the World Data Center for Remote Sensing of the Atmosphere (WDC-RSAT).

We acknowledge the use of NASA/GSFC’s Space Physics Data Facility’s OMNI dataset through OMNIWeb at <https://omniweb.gsfc.nasa.gov>, Papatashvili and King (2020), where we obtained the $F_{10.7\text{cm}}$ values used in this study. The QBO data have been obtained from Freie Universität (FU) Berlin: <https://www.geo.fu-berlin.de/en/met/ag/strat/produkte/qbo/index.html>.

Financial support. This research has been supported by the Bayerisches Staatsministerium für Umwelt und Verbraucherschutz (BayStMUV projects including GUDRUN, grant no. 71b-U8729-2003/125-13; GRIPS3 Back-Up, 2009/40051; BHEA, TLK01U-49580; LUDWIG, TUS01UFS-67093; VoCaS, TKP01KPB-70581; and AlpEn-DAC, TUS01UFS-72184).

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

Review statement. This paper was edited by Wen Yi and reviewed by Ravindra Pratap Singh, Pierre Simoneau, and two anonymous referees.

References

- Ammosov, P., Gavriljeva, G., Ammosova, A., and Koltovskoi, I.: Response of the mesopause temperatures to solar activity over Yakutia in 1993–2013, *Adv. Space Res.*, 54, 2518–2524, <https://doi.org/10.1016/j.asr.2014.06.007>, 2014.
- Baker, D. J. and Stair Jr., A. T.: Rocket measurements of the altitude distributions of the hydroxyl airglow, *Phys. Scripta*, 37, 611–622, <https://doi.org/10.1088/0031-8949/37/4/021>, 1988.
- Baldwin, M. P. and Dunkerton, T. J.: The solar cycle and stratosphere–troposphere dynamical coupling, *J. Atmos. Sol-Terr. Phys.*, 67, 71–82, <https://doi.org/10.1016/j.jastp.2004.07.018>, 2005.
- Batista, P. P., Takahashi, H., and Clemesha, B. R.: Solar cycle and the QBO effect on the mesospheric temperature and nightglow emissions at a low latitude station, *Adv. Space Res.*, 14, 221–224, [https://doi.org/10.1016/0273-1177\(94\)90139-2](https://doi.org/10.1016/0273-1177(94)90139-2), 1994.
- Beig, G., Keckhut, P., Lowe, R. P., Roble, R. G., Mlynczak, M. G., Scheer, J., Fomichev, V. I., Offermann, D., French, W. J. R., Shepherd, M. G., Semenov, A. I., Remsberg, E. E., She, C. Y., Lubken, F. J., Bremer, J., Clemesha, B. R., Stegman, J., Sigernes, F., and Fadnavis, S.: Review of mesospheric temperature trends, *Rev. Geophys.*, 41, 1015, <https://doi.org/10.1029/2002RG000121>, 2003.
- Bittner, M., Offermann, D., Bugaeva, I. V., Kokin, G. A., Koshelkov, J. P., Krivolutsky, A., Tarasenko, D. A., Gil-Ojeda, M., Hauchecorne, A., Lübken, F.-J., de la Morena, B. A., Mourier, A., Nakane, H., Oyama, K. I., Schmidlin, F. J., Soule, I., Thomas, L., and Tsuda, T.: Long period/large scale oscillations of temperature during the DYANA campaign, *J. Atmos. Terr. Phys.*, 56, 1675–1700, [https://doi.org/10.1016/0021-9169\(94\)90004-3](https://doi.org/10.1016/0021-9169(94)90004-3), 1994.
- Bittner, M., Offermann, D., and Graf, H. H.: Mesopause temperature variability above a midlatitude station in Europe, *J. Geophys. Res.-Atmos.*, 105, 2045–2058, <https://doi.org/10.1029/1999JD900307>, 2000.
- Bittner, M., Offermann, D., Graef, H. H., Donner, M., and Hamilton, K.: An 18-year time series of OH rotational temperatures and middle atmosphere decadal variations, *J. Atmos. Sol.-Terr. Phys.*, 64, 1147–1166, [https://doi.org/10.1016/S1364-6826\(02\)00065-2](https://doi.org/10.1016/S1364-6826(02)00065-2), 2002.
- Bittner, M., Höppner, K., Pilger, C., and Schmidt, C.: Mesopause temperature perturbations caused by infrasonic waves as a potential indicator for the detection of tsunamis and other geo-hazards, *Nat. Hazards Earth Syst. Sci.*, 10, 1431–1442, <https://doi.org/10.5194/nhess-10-1431-2010>, 2010.
- Brooke, J. S. A., Bernath, P. F., Wester, C. M., Sneden, C., Afşar, M., Li, G., and Gordon, I. E.: Line strengths of rovibrational and rotational transitions in the $X^2\Pi$ ground

- state of OH, *J. Quant. Spectrosc. Ra.*, 168, 142–157, <https://doi.org/10.1016/j.jqsrt.2015.07.021>, 2016.
- Dalin, P., Perminov, V., Pertsev, N., and Romejko, V.: Updated long-term trends in mesopause temperature, airglow emissions, and noctilucent clouds, *J. Geophys. Res.-Atmos.*, 125, e2019JD030814, <https://doi.org/10.1029/2019JD030814>, 2020.
- Espy, P. J., Ochoa Fernández, S., Forkman, P., Murtagh, D., and Stegman, J.: The role of the QBO in the inter-hemispheric coupling of summer mesospheric temperatures, *Atmos. Chem. Phys.*, 11, 495–502, <https://doi.org/10.5194/acp-11-495-2011>, 2011.
- French, W. J. R. and Klekociuk, A. R.: Long-term trends in Antarctic winter hydroxyl temperatures, *J. Geophys. Res.*, 116, D00P09, <https://doi.org/10.1029/2011JD015731>, 2011.
- French, W. J. R. and Mulligan, F. J.: Stability of temperatures from TIMED/SABER v1.07 (2002–2009) and Aura/MLS v2.2 (2004–2009) compared with OH(6–2) temperatures observed at Davis Station, Antarctica, *Atmos. Chem. Phys.*, 10, 11439–11446, <https://doi.org/10.5194/acp-10-11439-2010>, 2010.
- French, W. J. R., Burns, G. B., Finlayson, K., Greet, P. A., Lowe, R. P., and Williams, P. F. B.: Hydroxyl (6–2) airglow emission intensity ratios for rotational temperature determination, *Ann. Geophys.*, 18, 1293–1303, <https://doi.org/10.1007/s00585-000-1293-2>, 2000.
- French, W. J. R., Mulligan, F. J., and Klekociuk, A. R.: Analysis of 24 years of mesopause region OH rotational temperature observations at Davis, Antarctica – Part 1: long-term trends, *Atmos. Chem. Phys.*, 20, 6379–6394, <https://doi.org/10.5194/acp-20-6379-2020>, 2020.
- Goldman, A., Schoenfeld, W. G., Goorvitch, D., Chackerian Jr., C., Dothe, H., Mélen, F., Abrams, M. C., and Selby, J. E. A.: Updated line parameters for OH $X^2\Pi-X^2\Pi(v'', v')$ Transitions, *J. Quant. Spectrosc. Ra.*, 59, 453–469, [https://doi.org/10.1016/S0022-4073\(97\)00112-X](https://doi.org/10.1016/S0022-4073(97)00112-X), 1998.
- Gordon, I. E., Rothman, L. S., Hill, C., Kochanov, R. V., Tan, Y., Bernath, P. F., Birk, M., Boudon, V., Campargue, A., Chance, K. V., Drouin, B. J., Flaud, J.-M., Gamache, R. R., Hodges, J. T., Jacquemart, D., Perevalov, V. I., Perrin, A., Shine, K. P., Smith, M.-A. H., Tennyson, J., Toon, G. C., Tran, H., Tyuterev, V. G., Barbe, A., Csaszar, A. G., Devi, V. M., Furtenbacher, T., Harrison, J. J., Hartmann, J.-M., Jolly, A., Johnson, T. J., Karman, T., Kleiner, I., Kyuberis, A. A., Loos, J., Lyulin, O. M., Massie, S. T., Mikhailenko, S. N., Moazzen-Ahmadi, N., Muller, H. S. P., Naumenko, O. V., Nikitin, A. V., Polyansky, O. L., Rey, M., Rotger, M., Sharpe, S. W., Sung, K., Starikova, E., Tashkun, S. A., Vander Auwera, J., Wagner, G., Wilzewski, J., Weislo, P., Yu, S., and Zak, E. J.: The HITRAN2016 Molecular Spectroscopic Database, *J. Quant. Spectrosc. Ra.*, 203, 3–69, <https://doi.org/10.1016/j.jqsrt.2017.06.038>, 2017.
- Holmen, S. E., Dyrland, M. E., and Sigernes, F.: Long-term trends and the effect of solar cycle variations on mesospheric winter temperatures over Longyearbyen, Svalbard (78°N), *J. Geophys. Res.-Atmos.*, 119, 6596–6608, <https://doi.org/10.1002/2013JD021195>, 2014.
- Holtzclaw, K. W., Person, J. C., and Green, B. D.: Einstein coefficients for emission from high rotational states of the OH($X^2\Pi$) radical, *J. Quant. Spectrosc. Ra.*, 49, 223–235, [https://doi.org/10.1016/0022-4073\(93\)90084-U](https://doi.org/10.1016/0022-4073(93)90084-U), 1993.
- Höppner, K. and Bittner, M.: Evidence for solar signals in the mesopause temperature variability?, *J. Atmos. Sol.-Terr. Phys.*, 69, 431–448, <https://doi.org/10.1016/j.jastp.2006.10.007>, 2007.
- Kalicsinsky, C., Knieling, P., Koppmann, R., Offermann, D., Steinbrecht, W., and Wintel, J.: Long-term dynamics of OH* temperatures over central Europe: trends and solar correlations, *Atmos. Chem. Phys.*, 16, 15033–15047, <https://doi.org/10.5194/acp-16-15033-2016>, 2016.
- Kalicsinsky, C., Peters, D. H. W., Entzian, G., Knieling, P., and Mathias, V.: Observational evidence for a quasi-bidecadal oscillation in the summer mesopause region over Western Europe, *J. Atmos. Sol.-Terr. Phys.*, 178, 7–16, <https://doi.org/10.1016/j.jastp.2018.05.008>, 2018.
- Krassovsky, V. I., Shefov, N. N., and Yarin, V. I.: Atlas of the airglow spectrum 3000–12400 Å, *Planet. Space Sci.*, 9, 883, [https://doi.org/10.1016/0032-0633\(62\)90008-9](https://doi.org/10.1016/0032-0633(62)90008-9), 1962.
- Kvifte, G. I.: Nightglow observations at Ås during the I.G.Y., *Geophysica Norvegica*, 20, 1–15, 1959.
- Labitzke, K.: Sunspots, the QBO, and the stratospheric temperature in the north polar region, *Geophys. Res. Lett.*, 14, 535–537, <https://doi.org/10.1029/GL014i005p00535>, 1987.
- Labitzke, K.: On the solar cycle-QBO relationship: a summary, *J. Atmos. Sol.-Terr. Phys.*, 67, 45–54, <https://doi.org/10.1016/j.jastp.2004.07.016>, 2005.
- Lange, G.: Messung der Infrarotemissionen von OH* und O₂ in der Mesosphäre, PhD thesis, University of Wuppertal, WUB-DI 82-3, Wuppertal, 1982.
- Langhoff, S. R., Werner, H.-J., and Rosmus, P.: Theoretical transition probabilities for the OH Meinel system, *J. Mol. Spectrosc.*, 118, 507–529, [https://doi.org/10.1016/0022-2852\(86\)90186-4](https://doi.org/10.1016/0022-2852(86)90186-4), 1986.
- Laštovička, J. and Jelínek, Š.: Problems in calculating long-term trends in the upper atmosphere, *J. Atmos. Sol.-Terr. Phys.*, 189, 80–86, <https://doi.org/10.1016/j.jastp.2019.04.011>, 2019.
- Laštovička, J., Akmaev, R. A., Beig, G., Bremer, J., and Emmert, J. T.: Global change in the upper atmosphere, *Science*, 314, 1253–1254, <https://doi.org/10.1126/science.1135134>, 2006.
- Laštovička, J., Akmaev, R. A., Beig, G., Bremer, J., Emmert, J. T., Jacobi, C., Jarvis, M. J., Nedoluha, G., Portnyagin, Yu. I., and Ulich, T.: Emerging pattern of global change in the upper atmosphere and ionosphere, *Ann. Geophys.*, 26, 1255–1268, <https://doi.org/10.5194/angeo-26-1255-2008>, 2008.
- Lednyts'kyy, O., von Savigny, C., and Weber, M.: Sensitivity of equatorial atomic oxygen in the MLT region to the 11-year and 27-day solar cycles, *J. Atmos. Sol.-Terr. Phys.*, 162, 136–150, <https://doi.org/10.1016/j.jastp.2016.11.003>, 2017.
- Liu, W., Xu, J., Smith, A. K., and Yuan, W.: Comparison of rotational temperature derived from ground-based OH airglow observations with TIMED/SABER to evaluate the Einstein coefficients, *J. Geophys. Res.-Space*, 120, 10069–10082, <https://doi.org/10.1002/2015JA021886>, 2015.
- Meinel, A. B.: OH emission bands in the spectrum of the night sky. II, *Astrophys. J.*, 112, 120–130, <https://doi.org/10.1086/145321>, 1950.
- Mies, F. H.: Calculated Vibrational Transition Probabilities of OH($X^2\Pi$), *J. Mol. Spectrosc.*, 53, 150–188, [https://doi.org/10.1016/0022-2852\(74\)90125-8](https://doi.org/10.1016/0022-2852(74)90125-8), 1974.
- Nelson Jr., D. D., Schiffman, A., Nesbitt, D. J., Orlando, J. J., and Burkholder, J. B.: H + O₃ Fourier-transform in-

- frared emission and laser absorption studies of OH($X^2\Pi$) radical – An experimental dipole moment function and state-to-state Einstein A coefficients, *J. Chem. Phys.*, 93, 7003–7019, <https://doi.org/10.1063/1.459476>, 1990.
- Neumann, A.: QBO and solar activity effects on temperatures in the mesopause region, *J. Atmos. Terr. Phys.*, 52, 165–173, [https://doi.org/10.1016/0021-9169\(90\)90120-C](https://doi.org/10.1016/0021-9169(90)90120-C), 1990.
- Newman, P. A., Coy, L., Pawson, S., and Lait, L. R.: The anomalous change in the QBO in 2015–2016, *Geophys. Res. Lett.*, 43, 8791–8797, <https://doi.org/10.1002/2016GL070373>, 2016.
- Nikolashkin, S. V., Ignatyev, V. M., and Yugov, V. A.: Solar activity and QBO influence on the temperature regime of the subauroral middle atmosphere, *J. Atmos. Sol.-Terr. Phys.*, 63, 853–858, [https://doi.org/10.1016/S1364-6826\(00\)00207-8](https://doi.org/10.1016/S1364-6826(00)00207-8), 2001.
- Noll, S., Kimeswenger, S., Proxauf, B., Unterguggenberger, S., Kausch, W., and Jones, A. M.: 15 years of VLT/UVES OH intensities and temperatures in comparison with TIMED/SABER data, *J. Atmos. Sol.-Terr. Phys.*, 163, 54–69, <https://doi.org/10.1016/j.jastp.2017.05.012>, 2017.
- Noll, S., Winkler, H., Goussev, O., and Proxauf, B.: OH level populations and accuracies of Einstein-A coefficients from hundreds of measured lines, *Atmos. Chem. Phys.*, 20, 5269–5292, <https://doi.org/10.5194/acp-20-5269-2020>, 2020.
- Offermann, D., Hoffmann, P., Knieling, P., Koppmann, R., Oberheide, J., and Steinbrecht, W.: Long-term trends and solar cycle variations of mesospheric temperature and dynamics, *J. Geophys. Res.*, 115, D18127, <https://doi.org/10.1029/2009JD013363>, 2010.
- Papitashvili, N. E. and King, J. H.: OMNI Daily Data, NASA Space Physics Data Facility [data set], <https://doi.org/10.48322/5fmx-hv56>, 2020.
- Pautet, P. D., Taylor, M. J., Pendleton, W. R., Zhao, Y., Yuan, T., Esplin, R., and McLain, D.: Advanced mesospheric temperature mapper for high-latitude airglow studies, *Appl. Optics*, 53, 5934–5943, <https://doi.org/10.1364/AO.53.005934>, 2014.
- Perminov, V. I., Semenov, A. I., Medvedeva, I. V., and Zhelenov, Yu. A.: Variability of mesopause temperature from the hydroxyl airglow observations over mid-latitude sites, Zvenigorod and Tory, Russia, *Adv. Space Res.*, 54, 2511–2517, <https://doi.org/10.1016/j.asr.2014.01.027>, 2014.
- Perminov, V. I., Semenov, A. I., Pertsev, N. N., Medvedeva, I. V., Dalin, P. A., and Sukhodoev, V. A.: Multi-year behaviour of the midnight OH* temperature according to observations at Zvenigorod over 2000–2016, *Adv. Space Res.*, 61, 1901–1908, <https://doi.org/10.1016/j.asr.2017.07.020>, 2018.
- Pertsev, N. and Perminov, V.: Response of the mesopause airglow to solar activity inferred from measurements at Zvenigorod, Russia, *Ann. Geophys.*, 26, 1049–1056, <https://doi.org/10.5194/angeo-26-1049-2008>, 2008.
- Reisin, E. R. and Scheer, J.: Searching for trends in mesopause region airglow intensities and temperatures at El Leoncito, *Phys. Chem. Earth Pt. A/B/C*, 27, 563–569, [https://doi.org/10.1016/S1474-7065\(02\)00038-4](https://doi.org/10.1016/S1474-7065(02)00038-4), 2002.
- Reisin, E. R. and Scheer, J.: Unexpected East-West effect in mesopause region SABER temperatures over El Leoncito, *J. Atmos. Sol.-Terr. Phys.*, 157, 35–41, <https://doi.org/10.1016/j.jastp.2017.03.016>, 2017.
- Roble, R. G. and Dickinson, R. E.: How will changes in carbon dioxide and methane modify the mean structure of the mesosphere and thermosphere?, *Geophys. Res. Lett.*, 16, 1441–1444, <https://doi.org/10.1029/GL016i012p01441>, 1989.
- Roesch, A. and Schmidbauer, H.: WaveletComp: Computational Wavelet Analysis, R package version 1.1, CRAN [code], <https://CRAN.R-project.org/package=WaveletComp> (last access: 22 September 2023), 2018.
- Salby, M. and Callaghan, P.: Connection between the solar cycle and the QBO: The missing link, *J. Climate*, 13, 328–338, [https://doi.org/10.1175/1520-0442\(2000\)013<0328:CBTSCA>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<0328:CBTSCA>2.0.CO;2), 2000.
- Salby, M., Callaghan, P., and Shea, D.: Interdependence of the tropical and extratropical QBO: Relationship to the solar cycle versus a biennial oscillation in the stratosphere, *J. Geophys. Res.-Atmos.*, 102, 29789–29798, <https://doi.org/10.1029/97JD02606>, 1997.
- Scheer, J., Reisin, E. R., Espy, J. P., Bittner, M., Graef, H.-H., Offermann, D., Ammosov, P. P., and Ignatyev, V. M.: Large-scale structures in hydroxyl rotational temperatures during DYANA, *J. Atmos. Terr. Phys.*, 56, 1701–1715, [https://doi.org/10.1016/0021-9169\(94\)90005-1](https://doi.org/10.1016/0021-9169(94)90005-1), 1994.
- Schmidt, C.: Entwicklung eines bodengebundenen Infrarotspektrometers für die zeitlich hochaufgelöste Beobachtung des OH-Leuchtens aus der Mesopausenregion, PhD thesis, University of Augsburg, ISSN 1434-8454, <https://elib.dlr.de/108415/> (last access: 22 September 2023), 2016.
- Schmidt, C., Höppner, K., and Bittner, M.: A ground-based spectrometer equipped with an InGaAs array for routine observations of OH(3-1) rotational temperatures in the mesopause region, *J. Atmos. Sol.-Terr. Phys.*, 102, 125–139, <https://doi.org/10.1016/j.jastp.2013.05.001>, 2013.
- Schmidt, C., Küchelbacher, L., Wüst, S., and Bittner, M.: Nocturnal means of OH(3-1) airglow rotational temperatures (version 1.0) from the mesopause region obtained at the Environmental Research Station “Schneefernerhaus” between 2009 and 2020, Germany, WDC-RSAT [data set], <https://doi.org/10.26042/WDCRSAT.XZB5TZQG>, 2023a.
- Schmidt, C., Küchelbacher, L., Wüst, S., and Bittner, M.: Nocturnal means of OH(3-1) airglow rotational temperatures (version 1.0A) from the mesopause region obtained at the Environmental Research Station “Schneefernerhaus” between 2009 and 2020, Germany WDC-RSAT [data set], <https://doi.org/10.26042/WDCRSAT.Y0AEOEPZ>, 2023b.
- Semenov, A. I.: Long Term Temperature Trends for Different Seasons by Hydroxyl Emission, *Phys. Chem. Earth Pt. B*, 25, 525–529, [https://doi.org/10.1016/S1464-1909\(00\)00058-7](https://doi.org/10.1016/S1464-1909(00)00058-7), 2000.
- Shepherd, T. G., Koshyk, J. N., and Ngan, K.: On the nature of large-scale mixing in the stratosphere and mesosphere, *J. Geophys. Res.-Atmos.*, 105, 12433–12446, <https://doi.org/10.1029/2000JD900133>, 2000.
- Turnbull, D. N. and Lowe, R. P.: New hydroxyl transition probabilities and their importance in airglow studies, *Planet. Space Sci.*, 37, 723–738, [https://doi.org/10.1016/0032-0633\(89\)90042-1](https://doi.org/10.1016/0032-0633(89)90042-1), 1989.
- van der Loo, M. P. and Groenenboom, G. C.: Theoretical transition probabilities for the OH Meinel system, *J. Chem. Phys.*, 126, 114314, <https://doi.org/10.1063/1.2646859>, 2007.
- van der Loo, M. P. and Groenenboom, G. C.: Erratum: “Theoretical transition probabilities for the OH Meinel system” [J.

- Chem. Phys., 126, 114314 (2007)], J. Chem. Phys., 128, 159902, <https://doi.org/10.1063/1.2899016>, 2008.
- von Savigny, C., McDade, I. C., Eichmann, K.-U., and Burrows, J. P.: On the dependence of the OH* Meinel emission altitude on vibrational level: SCIAMACHY observations and model simulations, *Atmos. Chem. Phys.*, 12, 8813–8828, <https://doi.org/10.5194/acp-12-8813-2012>, 2012.
- Wendt, V., Wüst, S., Mlynczak, M. G., Russell III, J. M., Yee, J.-H., and Bittner, M.: Impact of atmospheric variability on validation of satellite-based temperature measurements, *J. Atmos. Sol.-Terr. Phys.*, 102, 252–260, <https://doi.org/10.1016/j.jastp.2013.05.022>, 2013.
- Wüst, S. and Bittner, M.: Non-linear resonant wave–wave interaction (triad): Case studies based on rocket data and first application to satellite data, *J. Atmos. Sol.-Terr. Phys.*, 68, 959–976, <https://doi.org/10.1016/j.jastp.2005.11.011>, 2006.
- Wüst, S., Wendt, V., Schmidt, C., Lichtenstern, S., Bittner, M., Yee, J. H., Mlynczak, M. G., and Russell III, J. M.: Derivation of gravity wave potential energy density from NDMC measurements, *J. Atmos. Sol.-Terr. Phys.*, 138–139, 32–46, <https://doi.org/10.1016/j.jastp.2015.12.003>, 2016.
- Wüst, S., Bittner, M., Yee, J.-H., Mlynczak, M. G., and Russell III, J. M.: Variability of the Brunt–Väisälä frequency at the OH* layer height, *Atmos. Meas. Tech.*, 10, 4895–4903, <https://doi.org/10.5194/amt-10-4895-2017>, 2017.
- Wüst, S., Bittner, M., Yee, J.-H., Mlynczak, M. G., and Russell III, J. M.: Variability of the Brunt–Väisälä frequency at the OH*-airglow layer height at low and midlatitudes, *Atmos. Meas. Tech.*, 13, 6067–6093, <https://doi.org/10.5194/amt-13-6067-2020>, 2020.
- Xu, J., Smith, A. K., Liu, H. L., Yuan, W., Wu, Q., Jiang, G., Mlynczak, M. G., Russell III, J. M., and Franke, S. J.: Seasonal and quasi-biennial variations in the migrating diurnal tide observed by Thermosphere, Ionosphere, Mesosphere, Energetics and Dynamics (TIMED), *J. Geophys. Res.-Atmos.*, 114, D13107, <https://doi.org/10.1029/2008JD011298>, 2009.