Intercomparison of stratospheric temperature profiles from a ground-based microwave radiometer with other techniques

Francisco Navas-Guzmán[1], Niklaus Kämpfer[1], Franziska Schranz[1], Wolfgang Steinbrecht[2], and Alexander Haefele[3]

[1]Institute of Applied Physics (IAP), University of Bern, Bern, Switzerland
[2]Meteorologisches Observatorium Hohenpeißenberg, Deutscher Wetterdienst, Hohenpeißenberg, Germany
[3]Federal Office of Meteorology and Climatology MeteoSwiss, Payerne, Switzerland

Correspondence to: Francisco Navas-Guzmán (francisco.navas@iap.unibe.ch)

Abstract.

In this work the stratospheric performance of a relatively new microwave temperature radiometer (TEMPERA) has been evaluated. With this goal in mind, almost three years of temperature measurements (January 2014 - September 2016) from the TEMPERA radiometer were intercompared with simultaneous measurements from other techniques: radiosondes, MLS satellite and Rayleigh lidar. This intercomparison campaign was carried out at the aerological station of MeteoSwiss at Payerne (Switzerland). In addition, the temperature profiles from TEMPERA were used to validate the temperature outputs from the SD-WACCM model. The results showed in general a very good agreement between TEMPERA and the different instruments and the model with a high correlation (higher than 0.9) in the temperature evolution at different altitudes between TEMPERA and the different datasets. An annual pattern was observed in the stratospheric temperature with generally higher temperatures in summer than in winter and with a higher variability during wintertime. A clear change in the tendency of the temperature deviations was detected in summer 2015 which was due to the repair of an attenuator in the TEMPERA spectrometer. The mean and the standard deviations of the temperature differences between TEMPERA and the different measurements were calculated for two periods (before and after the repair) in order to quantify the accuracy and precision of this radiometer over the campaign period. The results showed absolute biases and standard deviations lower than 2 K for most of the altitudes. And comparisons proved the good performance of TEMPERA in measuring the temperature in the stratosphere.

1 Introduction

The thermal structure of the atmosphere is one of the most important characteristics for determining chemical, dynamical and radiative processes in the atmosphere. In the stratosphere, temperature can influence chemical processes, and its vertical profile is fundamental to investigations of other atmospheric species, such as ozone and water vapor (Haefele et al., 2009; Stähli et al., 2013; Moreira et al., 2015). In addition, stratospheric temperature is a very important indicator of climate change (Randel et al., 2009). The temperature trends can provide evidence of the roles of natural and anthropogenic climate change mechanisms. Several studies have shown the observation of a pattern of tropospheric warming and lower stratospheric cooling during the last few decades of the twentieth century which is very likely related to anthropogenic emissions of trace gases,

ozone and aerosols (Ramaswamy and Schwarzkopf, 2002; Santer et al., 2006; Schwarzkopf and Ramaswamy, 2008; Randel et al., 2009; Bindoff et al., 2013).

Stratospheric temperatures can present a large variability in time, specially during winter. For example, the stratosphere can experience sudden temperature increases (Sudden Stratosphere Warming, SSW) due to dynamical processes where the temperature can change by several tens of degrees within a very short time (Flury et al., 2009; Scheiben et al., 2012). Monitoring these fast changes require measurement techniques with high temporal and spatial resolution.

The in situ technique of radiosonde is extensively used for tropospheric temperature measurements due to its high vertical resolution. However, radiosondes are only able to cover the lower part of the stratosphere, reaching a maximum altitude of around 35 km. In addition, since at best they are launched four times a day, they offer only a very low temporal resolution compared with other techniques.

At present, stratospheric temperature profiles are mostly obtained by remote sensing methods, such as lidars and microwave radiometers. Rayleigh lidars have been shown to be a powerful tool for monitoring temperatures in the middle atmosphere with a high spatial and temporal resolution (Hauchecorne and Chanin, 1980; Keckhut et al., 2001; Steinbrecht et al., 2009). However this technique's main drawback is that they cannot be operated during daytime, or under cloudy or rainy conditions. Microwave radiometer measurements can overcome these difficulties, since the measurements in the microwave region are almost unaffected by liquid water and the radiometers can be continuously operated providing temperature profiles with a reasonably good spatial and temporal resolution. Most of the microwave radiometers for stratospheric temperature measurements are operated on board satelllites (e.g. MLS instrument on the Aura satellite as described in Waters et al. (2006), AMSU-A instrument on the Aqua satellite as described in Aumann et al. (2003) and the SABER instrument on the TIMED satellite as described in Remsberg et al. (2003)).

The possibility of using ground-based microwave radiometry for stratospheric temperature measurements was first shown in Waters (1973) and it has recently been implemented (Shvetsov et al., 2010; Stähli et al., 2013). The technique is based on the stratospheric thermal emission from high-rotational magnetic dipole transitions of molecular oxygen around 53 GHz. The main advantages of ground-based radiometer measurements are that they can provide unattended continuous measurements of temperature profiles in almost all weather conditions with reasonably good spatial and temporal resolution in the altitude range between 20 and 50 km above sea level (asl). In addition, long-term measurements in a fixed location allow the local atmospheric thermodynamics to be characterized. In this study we are going to present almost three years of stratospheric temperature measurements from the TEMPErature RAdiometer (TEMPERA) which has been designed and built by the Institute of Applied Physics of the University of Bern (Switzerland). This is the first ground-based microwave radiometer that is able to retrieve temperature measurements in the troposphere and in the stratosphere at the same time. Tropospheric retrievals from this radiometer have been evaluated in detail in other studies (Stähli et al., 2013; Navas-Guzmán et al., 2014, 2016). In this work we will focus on the stratospheric performance of TEMPERA (from 20 to 50 km) comparing its measurements with the ones from different instruments and techniques: radiosondes, satellite and lidar measurements. In addition TEMPERA profiles will be used to validate the temperature outputs from SD-WACCM model.

The results obtained in this study provide a detailed evaluation of the temperature retrievals from the TEMPERA radiometer. The paper has been organized in the following way. The description of the different instrumentation used in this work is introduced in Section 2. Section 3 presents a detailed description of the methodology used for the microwave temperature retrievals. Section 4 presents the results of the different comparisons of RS, MLS satellite, lidar, SD-WACCM versus the TEMPERA radiometer. And finally, we conclude with a summary of the key findings in Section 5.

2   Experimental site and instrumentation

A special campaign has been set up at the aerological station in Payerne (46.82° N, 6.95° E; 491 m above sea level (asl), Switzerland) of the Swiss Federal Institute of Meteorology and Climatology (MeteoSwiss). For this campaign, the TEMPERA radiometer was moved from the ExWi building of the University of Bern (Bern, Switzerland) to Payerne in December 2013. The main goal of this campaign is to assess the tropospheric and stratospheric performance of TEMPERA using the versatile instrumentation available at this MeteoSwiss station (Navas-Guzmán et al., 2016). In particular, this study will focus on the intercomparison of the stratospheric temperature profiles from TEMPERA.

Next, we will introduce the ground-based microwave radiometer called TEMPERA and all the other instrumentation used in this study. As already mentioned, the TEMPERA radiometer is the first ground-based microwave radiometer able to measure temperature profiles in the troposphere and in the stratosphere simultaneously (Stähli et al., 2013; Navas-Guzmán et al., 2014, 2016). It measures the microwave emission of the molecular oxygen in the 51-57 GHz range. The instrument consists of a frontend to collect the microwave radiation and two backends for the spectral analysis (a filter bank and a Fast Fourier Transform (FFT) spectrometer). The incoming radiation is directed into a corrugated horn antenna using an off-axis parabolic mirror. The antenna is characterized by a Half Power Beam Width (HPBW) of 4°. The detected signal in the two backends is calibrated by means of an ambient hot load in combination with a noise diode. The calibration of the noise diode is performed every month using a hot (ambient) and a cold (liquid nitrogen) load. Figure 1 (left) shows a picture of TEMPERA radiometer where its different components can be observed: mirror (1), microwave absorbers (hot(2) and cold (3) load), receiver (4) and styrofoam window (5). Figure 1 (right) shows the isolated room where TEMPERA is located at the MeteoSwiss aerological station in Payerne (Switzerland).

The tropospheric measurements by TEMPERA are performed by means of a filter bank. It covers a total of 12 frequencies uniformly distributed on the wing of the 60 GHz oxygen emission complex. Since tropospheric temperature measurements are not the topic of this study more details about technical aspects of the filter bank and the measurement protocol for this mode can be found in Stähli et al. (2013) and Navas-Guzmán et al. (2016).

For stratospheric measurements a second backend is used. It consists of a digital FFT spectrometer (Acqiris AC240) which measures the two pressure-broadened oxygen emission lines centered at 52.5424 and 53.0669 GHz. The bandwidth of this spectrometer is 960 MHz and has a resolution of 30.5 kHz. The receiver noise temperature T_N is around 480 K. More technical details about the different components of the microwave receiver, such as the IQ-Mixer and the local oscilator (LO), can be

Figure 1. The TEMPERA instrument at the MeteoSwiss Station in Payerne, Switzerland.
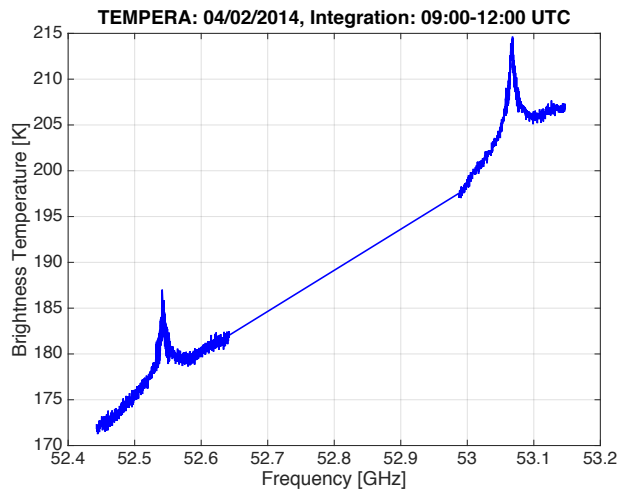
Figure 2. Spectrum of brightness temperatures measured with TEMPERA on 4 February 2014 from 09:00-12:00 (UTC). Only the FFT channels of the the the first line at 52.5424 GHz and the second line at 53.0669 GHz used in the temperature retrievals are shown.

found in Stähli et al. (2013). An example of a calibrated spectrum (brightness temperature) measured with this spectrometer on 2 of February of 2014 is shown in Fig. 2.

A styrofoam window allows views of the atmosphere over a range of different elevation angles (from $20°$ to $60°$). The radiometer is operated inside a laboratory primarily to protect it against adverse weather conditions. The frontend has additional

temperature stabilization using Peltier elements in combination with a ventilation system that allows the frontend plate to be stabilized to within ±0.2 K (Stähli et al., 2013).

Every measurement cycle takes 1 minute and starts with a calibration using the hot load in combination with a noise diode for 9 s, followed by atmosphere measurements. These atmospheric measurements consist of scanning from $20°$ to $60°$ elevation in steps of $5°$ (9 angles). The observations at all the angles are used for tropospheric measurements while only the observations at $60°$ elevation angle, which take 15 seconds, are used for stratospheric measurements (Stähli et al., 2013). Details about the methodology used to obtain stratospheric temperature profiles from these measurements will be given in section 3.

Independent in-situ temperature measurements have been taken by means of radiosondes. They have been launched twice a day at the aerological station of Payerne since 1954. The target level of radiosondes is 10 hPa (approx. 32 km), and hence cover only the lower stratosphere. Their spatial resolution ranges between 10 and 80 m with a highest resolution in the first seconds of the flight. The Swiss Radiosonde SRS-C34 introduced in 2011 uses a thermocouple for temperature measurements and a polymer hygristor for relative humidity measurements. Pressure is calculated from temperature and GPS altitude assuming hydrostatic equilibrium. The achieved uncertainties are $\pm$ 0.2 K for temperature, $\pm$ 2 hPa (accuracy increases with height) for pressure and $\pm$ 5 to 10% for relative humidity.

Stratospheric temperature have been also obtained from the Microwave Limb Sounder (MLS) instrument on board of the Aura satellite. MLS has been making measurements of atmospheric composition, temperature, humidity and cloud ice in the upper troposphere, stratosphere and lower mesosphere since August 2004 (Waters et al., 2006). It observes thermal microwave emission from Earth's limb viewing forward along the Aura spacecraft flight direction, scanning its view from the ground to 90 km every 25 seconds. Aura is in a near-polar 705 km altitude orbit. As Earth rotates underneath it, the Aura orbit stays fixed relative to the sun; to give daily global coverage with 15 orbits per day. Aura is part of NASA's A-train group of Earth observing satellites. These satellites fly in formation with the different satellites making measurements within a short time of each other. Temperature profiles are retrieved from MLS measurements using radiances near the O_2 spectral bands at 118 GHz for the stratosphere and mesosphere and at 239 GHz for the troposphere (Yan et al., 2016) using the optimal estimation theory (Rodgers, 2000). Four different versions of MLS data have been released to date. The initial version 1.5 (v1.5), was replaced by version 2.2/2.3 (v2) in 2007 and version 3.3/3.4 (v3) in 2010. The most recent production version, version 4.2 (v4), replaced v3 in February 2015. All the MLS data presented in this study correspond to the latest version (v4).

Temperature measurements in the upper stratosphere have been also obtained from a lidar at Hohenpeißenberg, Germany ($47.8°$ N, $11.0°$ E). This lidar has been operated since September 1987 by the German Weather Service (DWD) and has provided one of the longer NDACC time series (Steinbrecht et al., 2009). It emits intense ultraviolet light pulses at 353 nm generated from a Xenon Chloride excimer laser and a Hydrogen Raman cell. Light intensity scattered back from air molecules in the atmosphere (by Rayleigh scattering) is recorded as a function of altitude (=time from pulse emission to reception of backscattered light). Above the stratospheric aerosol layer, that is above 25 to 30 km, the returned light intensity is proportional to air density. Assuming hydrostatic equilibrium, this (relative) density profile can be integrated downward over altitude, providing a (relative) pressure profile. Division of the (relative) pressure profile by the (relative) density profile then yields the temperature profile. See Hauchecorne and Chanin (1980) for details. The method requires an initial guess for temperature (or
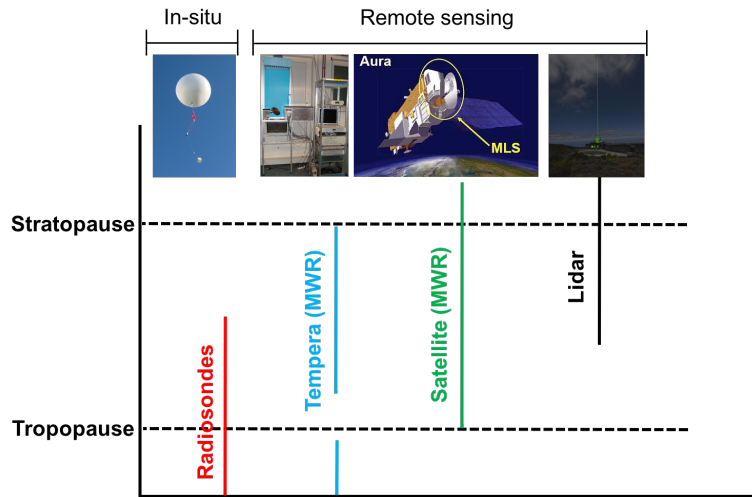
Figure 3. Measurement ranges for the different techniques used in this study (radiosondes, Tempera radiometer, MLS and lidar).

pressure) at the far end around 70 to 80 km altitude, but because of the large increase of pressure with decreasing altitude, this choice of initial value has virtually no influence on the derived temperatures below around 50 to 60 km altitude. The lidar requires clear nights for operation, and typically provides 80 to 90 nightly mean temperature profiles per year. The precision of the derived temperature is about ± 0.5 K at 30 km, ± 1 K at 45 km, ± 5 K at 60 km and ± 10 K at 70 km (all 1 sigma). Vertical

5 resolution is about 1.5 km. The lidar derived temperature has a small bias of about 2 K between 30 and 50 km, which is not well understood. See Steinbrecht et al. (2009) for details.

Figure 3 shows different ranges of measurements of each instrument used in this study (radiosondes, TEMPERA radiometer, MLS satellite and lidar). As we can see TEMPERA is the only instruments which is able to cover almost the full troposphere and stratosphere.

10 ## 3 Methodology

3.1 Temperature profiles from TEMPERA radiometer

Oxygen is a well-mixed gas whose fractional concentration is independent of altitude below approx. 80 km, so the microwave radiation from it contains information primarily on atmospheric temperature. The retrievals of stratospheric temperature profiles from TEMPERA are based on the measurements of two oxygen emission lines centered at 52.54 and 53.06 GHz (see Fig. 2).

15 The shape of these lines is governed by a pressure broadening mechanism up to 60 km of altitude, therefore the measured spectra can provide vertical information. The wings of the emission lines provide information of the radiation coming from low altitudes (higher broadening caused by higher pressure) while the center of the lines give information of the radiation

coming from upper altitudes (smaller broadening and lower pressure). Both emission lines measured by TEMPERA are used at the same time with a bandwidth of 200 MHz around the first line and of 160 MHz around the second. Only measurements at the highest elevation angle ($60°$) are used for stratospheric measurements with the digital FFT spectrometer. This limits the integration time with the FFT spectrometer to 15 s in each minute measurement cycle. In order to get a low enough noise level the measurements are integrated for half an hour which requires two hours of measurement time, since only one quarter of the measurement time is used for the digital FFT spectrometer (Stähli et al., 2013). Therefore the time resolution of the stratospheric temperature profiles from the TEMPERA radiometer is two hours.

Obtaining temperature profiles from the calibrated brightness temperature spectrum, an example of which is shown in Fig. 2, requires a solution to the radiative transfer equation. A unique solution does not exist, so some statistical constraints are needed in order to obtain physically meaningful solutions. In our case we use the optimal estimation method (OEM) (Rodgers, 2000) by means of the radiative transfer model ARTS/QPack (Eriksson et al., 2011). The method is based on Bayes' probability theorem and a detailed description of its application to TEMPERA measurements can be found in Stähli et al. (2013).

The ARTS package implements the radiative transfer equation (forward model), simulating the brightness temperature as:

$$\mathbf{y} = F(\mathbf{x}, \mathbf{b}) + \epsilon \tag{1}$$

where F denotes the forward model, the vector $\mathbf{y}$ corresponds to the measured spectrum (brightness temperature), $\mathbf{x}$ is the true temperature profile, $\mathbf{b}$ contains some additional forward model parameters, and ϵ is the measurement noise.

The solution to the inverse problem is obtained by using the Gauss-Newton iterative method, whose solution can be expressed in a matrix notation as follows:

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \left(\mathbf{S}_a^{-1} + \mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} \mathbf{K}_i^{-1} \right) \left[\mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} \left(\mathbf{y} - F(\mathbf{x}_i) \right) - \mathbf{S}_a^{-1} (\mathbf{x}_i - \mathbf{x}_a) \right] \tag{2}$$

where the vector $\mathbf{x}$ is the true temperature profile, $\mathbf{y}$ is the measured spectrum (brightness temperature), $\mathbf{x}_a$ is the a priori temperature profile, $\mathbf{S}_a$ is the a priori covariance matrix and $\mathbf{S}_\epsilon$ is the observation error-covariance matrix. The forward model is denoted by F, and the vector $\mathbf{K}$ is the weighting function ($\mathbf{K} = \partial F / \partial \mathbf{x}$).

An important tool used very often in the OEM is the averaging kernel matrix $\mathbf{A}$ (Rodgers, 2000). This matrix describes the response of the retrieved temperature profile $\hat{\mathbf{x}}$ to the true temperature profile $\mathbf{x}$ and is defined as:

$$\mathbf{A} = \mathbf{D}_y \mathbf{K}_x = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{x}} \tag{3}$$

where $\mathbf{K}_x$ is the weighting function already defined, and $\mathbf{D}_y = \partial F / \partial \mathbf{x}$ is the so-called contribution function.

The rows of $\mathbf{A}$ are called the averaging kernels (AVK) and they describe the sensitivity of the retrieval for a certain height level to a perturbation at other levels. The sum of the AVK is called the measurement response (MR), which describes the contribution of measurements to the retrieved profile at a certain height.

The method needs an a priori temperature profile in order to constrain the solutions to physically meaningful results. As a priori profiles, monthly mean temperature profiles from radiosonde measurements at Payerne from 1994 to 2011 are used in the lower part (ground to 15 km) and mean MLS temperature profiles from a climatology are used in the upper part. As a priori

covariance matrix $\mathbf{S}_a$ a function decreasing exponentially with a correlation of 3 km is used assuming a standard deviation of 2 K. For the observation errors the residuals of the inversion are considered (difference between the integrated spectra and the fit of the spectra). Under regular conditions these errors range between 0.5 and 1.5 K (Stähli et al., 2013).

In the radiative transfer calculations ($F(\mathbf{x}, \mathbf{b})$) the absorption coefficients of the different species are calculated using different
5   models: Rosenkranz (1998) for H_2O, Rosenkranz (1993) for O_2 and Liebe et al. (1993) for N_2. The density profiles of oxygen (O_2) and nitrogen (N_2) are incorporated by ARTS assuming standard atmospheric profiles for summer and winter (Anderson et al., 1986). In the case of tropospheric water vapor a profile with an exponential decrease is considered. This profile is calculated with the measured surface water vapor density from a weather station and assuming a scale height of 2000 m (Bleisch et al., 2011).

10   Figure 4 shows an example of temperature inversion from TEMPERA measurements using the OEM result obtained on 1 October 2015 for the time interval from 22 to 00 UTC. In Fig. 4a we can observe that the forward model brightness temperatures (red lines) agree well with the measured brightness temperatures (black lines), except for around the line center. The larger differences observed in the center of the emission lines (see Fig. 4b), is mainly due to a different binning used in the center of the lines and on the wings of the lines (Stähli et al., 2013). In addition, the Zeeman effect could explain some small differences
15   in the center of the lines since it is not incorporated in the forward model (Navas-Guzmán et al., 2015). Fig. 4c presents the a priori temperature profile used in the inversion (black dashed line) and the retrieved temperature profile (blue line). Figure 4d shows the averaging kernels (black lines), the measurement response (red line) and the height resolution, which is defined as the full-width at half-maximum (FWHM) of the averaging kernels (blue line). We can observe that for this inversion the height resolution ranges between 13 and 16 km. The MR shows values larger than 0.8 in the range between 20 and 43 km, meaning that
20   80% of the contribution to the retrieved temperature profile comes from the measurements. These values decrease with altitude reaching 0.5 at 47 km for this case. We would like to point out that the altitude range of the stratospheric temperatures from the TEMPERA radiometer used in this study correspond to levels with a high MR (higher than 0.8 at most of the altitudes). Finally, the total, observational (random error due to measurement noise) and smoothing errors are also calculated with this method and are shown in Fig. 4e.

25   In order to compare the temperature profiles from the different instruments (RS, MLS satellite, lidar) and also from the WACCM model with the ones from TEMPERA radiometer the profiles are first interpolated to the pressure grid of TEMPERA, and then are convolved with the averaging kernel of this radiometer in order to take into account the different height resolutions. Equation 4 gives the expression for calculating the convolved temperature profiles:

$$\hat{\mathbf{x}}_r = \mathbf{x}_a + \mathbf{A}\left(\mathbf{x}_r - \mathbf{x}_a\right) \tag{4}$$

30   where $\mathbf{x}_a$ is the a priori profile of the radiometer, $\mathbf{A}$ is the averaging kernel and $\mathbf{x}_r$ is the interpolated reference profile.

4   Results: Evaluation of stratospheric temperature profiles from TEMPERA

The TEMPERA radiometer has been almost continuously measuring since 2014 at the aerological station of MeteoSwiss at Payerne (Switzerland). Figure 5 (left) shows the stratospheric temperature evolution obtained from TEMPERA for the almost
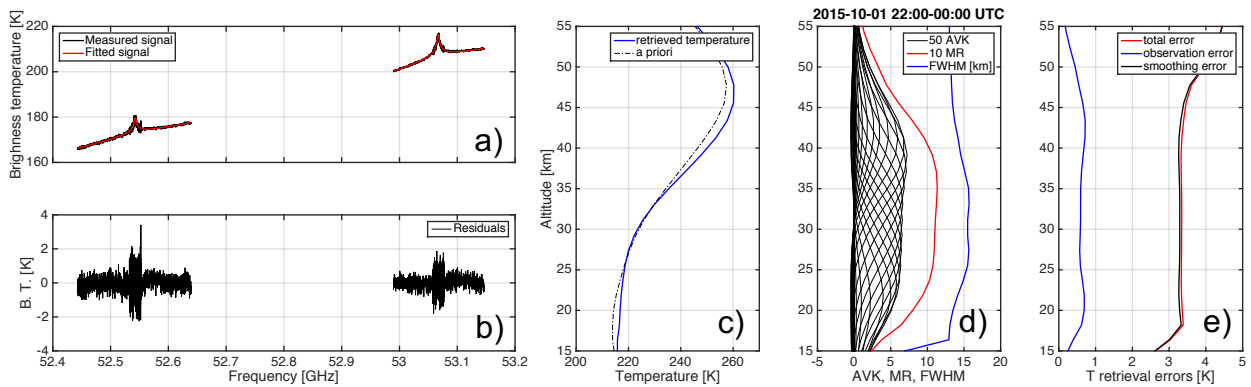
Figure 4. Temperature retrieval of 1 October 2015 using the optimal estimation method (OEM). a) Brightness temperature measured with TEMPERA (black lines) compared with the forward model brightness temperature (red lines) obtained for this retrieval. b) Residuals for this inversion. c) Retrieved temperature and a priori profile. d) Averaging kernels, measurement response and FWHM [km]. d) Temperature retrieval errors.

three years of measurements. From this plot a clear annual pattern can be observed with generally higher temperatures in spring and summer than in autumn and winter. Some interesting episodes can also be observed during the three presented winters, in which strong increases of temperature are measured for short periods in the upper stratosphere and could be identified as SSW. These increases in temperature in the upper stratosphere are often associated with a decrease in temperature in the lower stratosphere, which is a pattern characteristic of SSW events. Figure 5 (right) shows an example of strong variation of temperature in the stratosphere for a winter day (3 January 2015). In this case, the temperature changed up to 15 K for some altitudes in the course of only 10 hours. These measurements show the importance of continuous observations for a fixed location, since the important variations in temperature observed cannot be captured by only occasional measurements or measurements with poor temperature resolutions.

The temperature profiles from TEMPERA have been compared with those from other instruments and the SD-WACCM model, all of which have different spatial and temporal resolutions. Figure 6 presents three representative examples of stratospheric temperature profiles: one in winter, one in summer and one in autumn. Measurements from the different instruments and model (re)analysis show a generally good agreement in the range where they are comparable. Some differences are evident in the upper stratosphere between MLS measurements and the other profiles on 4 February 2014. For the other two days the lidar (black line) is the source that exhibits deviations with respect to the microwave measurements and the model in some ranges in the upper stratosphere. We point out the good agreement observed between TEMPERA radiometer and most of the other techniques in these three cases. The examples also illustrate the different vertical ranges and the spatial resolutions for the different measurements. We can observe that radiosondes only cover the lower stratosphere but with a high spatial resolution,
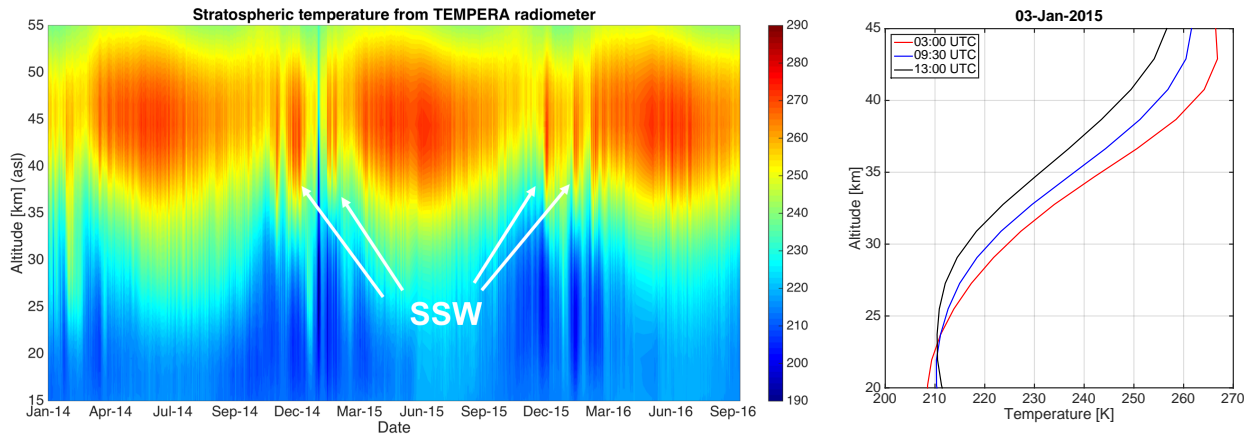
Figure 5. Left: Stratospheric temperature evolution from TEMPERA radiometer. Some SSW events are indicated by white arrows.
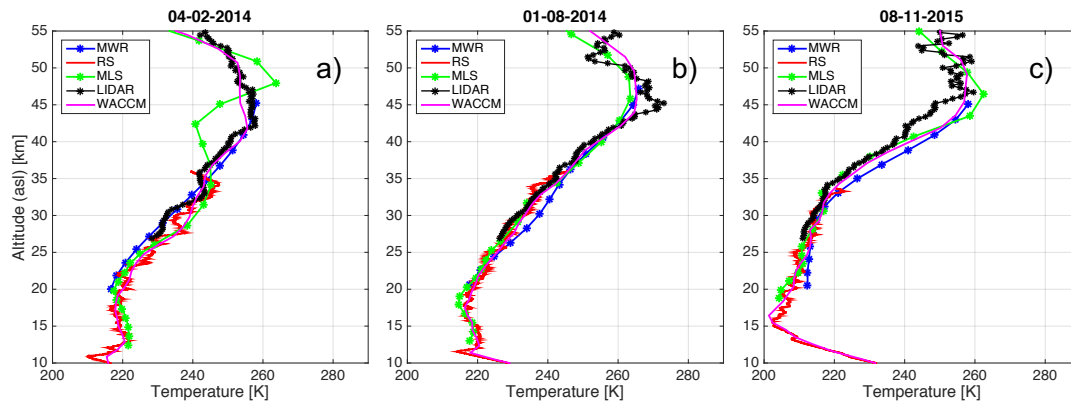
Figure 6. Stratospheric temperature profiles for night-time measurements from TEMPERA, RS, MLS, Lidar and WACCM model on (a) 4 February 2014, (b) 1 August 2014 and (c) 8 November 2015.

while lidar measurements provide information in the upper stratosphere. MLS and TEMPERA are able to cover almost the whole stratosphere although their spatial resolution is lower.

In order to validate the accuracy and errors of the temperature profiles from the TEMPERA radiometer a statistical analysis is performed with almost the three years of measurements. In sections 4.1, 4.2, 4.3 and 4.4 a comparison is made to, respectively,

5 RS measurements, MLS measurements, lidar measurements and the SD-WACCM model. A multiway comparison between all of these is then presented in section 4.5.

4.1 Comparison with RS

Stratospheric temperature profiles from TEMPERA have been compared with the ones from RS measurements for the period from January 2014 to September 2016. As indicated in previous sections radiosondes have been launched twice a day (11 and 23 UTC) at the aerological station at Payerne since 1954. The TEMPERA profiles closest in time to the RS launches have been selected in order to do this comparison. A total of 1489 pairs of profiles are used in these statistics, which were measured under all weather conditions except for rainy cases. The RS profiles were interpolated to the altitude grid of TEMPERA radiometer, and completed in the upper part with the TEMPERA measurements, since RSs usually do not reach altitudes higher than 30-35 km. Afterwards, the profiles were convolved using the averaging kernels of TEMPERA.

Figure 7 shows the temporal evolution of the stratospheric temperature at different altitudes from TEMPERA and RS for the campaign period. The interpolated temperatures from RS have also been plotted (green lines) in order to visualise the smoothing effect on them when they are convolved with the averaging kernels of TEMPERA. In addition, the a priori temperature used for the TEMPERA inversions is shown. The temperature deviations along this period between TEMPERA and the convolved measurements from RS are shown in the lower panels (black lines). We can observe in general a very good agreement between both instruments for the displayed altitudes with correlation coefficients higher than 0.9. An annual pattern is observed in the stratospheric temperature with higher temperatures in summer than in winter. Again in this plot we can observe that the variability of the temperature is higher during winter than in other seasons, and some interesting events with a strong increase in temperature have been detected (January 2014 and 2015, February 2016). The temperature deviations between TEMPERA and RS are in general small with most of the values below 3 K, although some short periods with larger discrepancies are also found (e.g. February 2015). We can also observe from these plots that the deviations at 27 km altitude are larger and noisier than for the other two altitudes. A remarkable feature observed in the temperature deviation lines at all the profiles is a small step in summer of 2015. This step is more evident in the two higher altitudes (27 km and 33 km) where the deviations changed from positive to negative. The effect is smaller at the lowest altitude (21.5 km) where it looks to have an opposite behaviour, changing from negative or almost zero deviations to positive deviations after the step happens. This change of tendency could be due to the fact that an attenuator in the FFT spectrometer was repaired in summer 2015. It seems that after this repair the brightness temperature spectra measured by the FFT were slightly affected and some small differences in the retrieved temperatures are observed.

In order to take into account this instrumental modification and characterize possible changes in the accuracy and precision of the TEMPERA radiometer the statistical analysis between TEMPERA and the other measurements (RS, MLS, lidar and WACCM) is carried out over two different measurement periods. From here on, period 1 will refer to the period before the attenuator in the FFT spectrometer was changed (January 2014-June 2015) and period 2 will refer to the period after this repair (July 2015-September 2016). In addition, the measurements have been split by season into winter and summer, with summer referring to April-September and winter October-March, inclusive. It is useful to make this distinction because there is a greater level of atmospheric variability in winter, which could produce larger deviations between the different measurements than those due to fundamental differences in the measurement techniques.
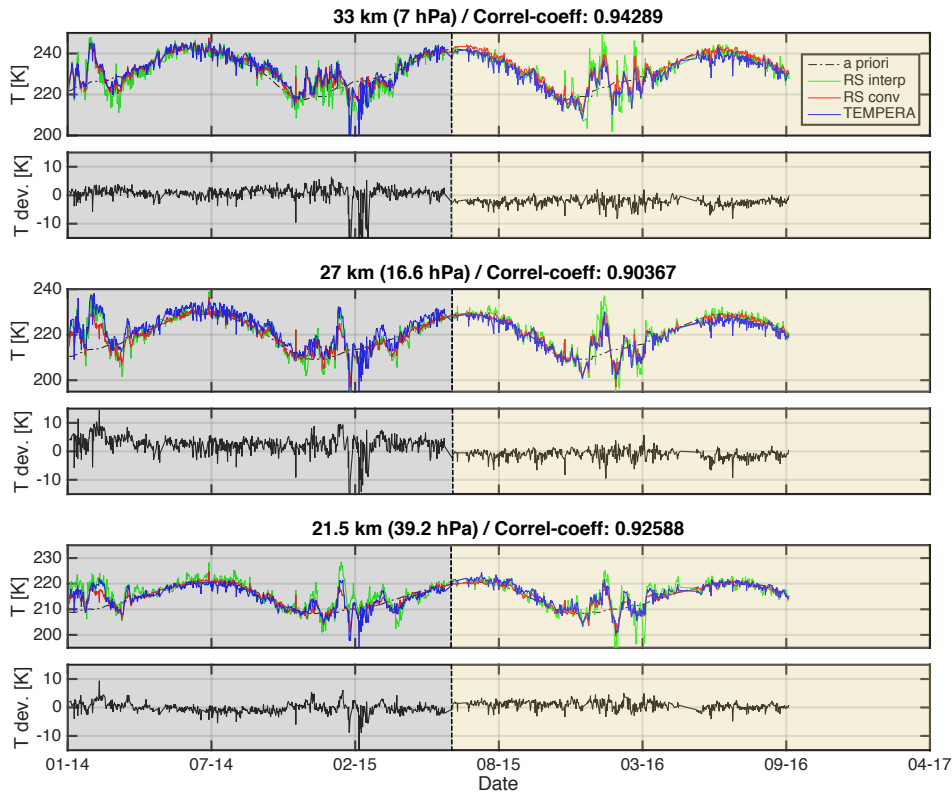
Figure 7. Stratospheric temperature evolution and temperature deviations at different altitudes for RS and TEMPERA. Different background colors are used to distinguish between period 1 and 2 (gray and light brown, respectively).

Figure 8 shows the mean and the standard deviations between TEMPERA and RS which have been calculated for all the measurements in each period (black lines) and also for winter and summer seasons of the different periods (blue and red lines, respectively). From this plot we can observe that there is a clear change in the mean bias between TEMPERA and RS for periods 1 and 2. The mean bias for period 1 ranged between -0.3 K at 20 km and 2.6 K at 28.5 km showing in general a positive

5 deviation at most of the altitudes. The mean bias in period 2 showed negative values for most of the altitudes with values ranging between 0.9 K (20 km) and -2.3 K (32 km). There is also a clear difference in the standard deviation observed for both periods. Period 1 showed much larger standard deviations than period 2 with values that range between 1.9 K (21 km) and 3.5 K (28.5 km). The standard deviations for period 2 were smaller and much more constant in height with values ranging between 1.3 K (34 km) and 1.7 K (26.5 km). These results show a change in the sign of the bias between TEMPERA and RS when the

10 attenuator of the FFT spectrometer was repaired in June of 2015, although in term of absolute values the differences were not very significant. However, the standard deviations for period 2 were smaller than for period 1 indicating a higher precision of the TEMPERA radiometer after the repair with respect to the reference RS measurements. If we have a look at the seasonal
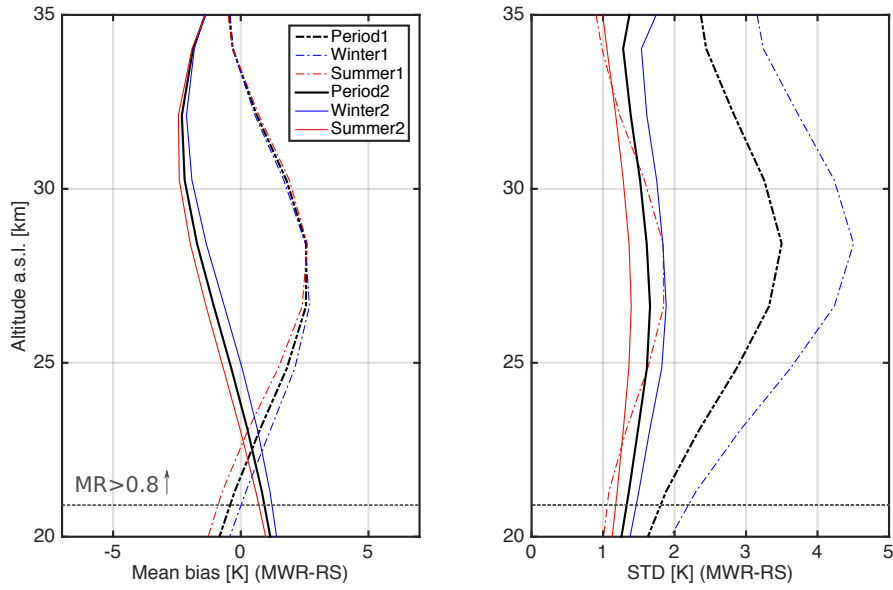
Figure 8. Mean temperature biases and standard deviations between TEMPERA and RS. A total of 1489 profiles have been compared (Period 1: 809 prof., dashed lines; Period 2: 680 prof., solid lines). The mean biases and the standard deviations for each periods are represented by black lines. The winter season is indicated with blue lines while the summer is indicated by red lines (Winter1: 421 prof.; Summer1: 388 prof.; Winter2: 289 prof. Summer2: 391).

behaviour of the bias for both periods we can observe that there are small differences between winter and summer. In the case of period 1 the maximum difference between winter and summer is 0.9 K and it is observed in the lower part, while for period 2 the differences are lower than 0.7 K. Much larger differences are found for the standard deviation between the two seasons for period 1 (dashed lines). While the standard deviations ranges between 0.9 K and 1.8 K in summer, the values ranged between 2
5 K and 4.5 K in winter, reaching the maximum standard deviation at 28.5 km. Although during period 2 the standard deviations in winter were also larger than in summer, the differences were not so remarkable (smaller than 0.5 K). These results show that there was a larger variability in the temperature deviations between TEMPERA and RS during the winters of period 1. It is something that could be expected from the temperature evolution showed in Fig. 7 that showed larger discrepancies specially during winter 2015.

10 **4.2 Comparison with Aura/MLS**

The stratospheric temperature profiles from TEMPERA have also been compared with those obtained from the MLS instrument on board the Aura satellite. As indicated in section 2 the temperature profiles used for MLS correspond to the version 4 retrievals. In order to select the temperature profiles from MLS to be used in the comparison we chose those that where collocated with the measurement site, which by our criteria meant that the MLS measurements were within $\pm 1°$ (± 110 km) of
15 the measuring site in latitude and $\pm 5°$ (± 460 km) in longitude. The data were also restricted to cases with near time-coincident

between TEMPERA and MLS, which means that the MLS profiles were taken during the period of the spectral integration for the TEMPERA measurements. A total of 367 profiles were obtained under these criteria and for all weather conditions excluding rainy cases. The MLS temperature profiles were interpolated to the pressure grid of TEMPERA and these profiles were convolved using the averaging kernels of TEMPERA as described in section 3.

5 Figure 9 shows the evolution of the stratospheric temperatures and the deviations between TEMPERA and MLS at 3 different altitude levels. Similar patterns to those observed in Fig. 7 are found in this plot (although with less data), observing an annual cycle with higher temperatures in summer than in winter and with a larger variability during wintertime. We can observe from these plots a very good agreement between both instruments despite the very different type of observations that we are comparing (ground-based against satellite measurements). This good agreement is also observed when strong variations in

10 temperature occur in a short time interval, as can be seen in the winter of 2016, and is confirmed by the high correlation coefficient (larger than 0.92) found at the different altitudes. The temperature deviations (TEMPERA-MLS) observed are in general small, although we observe some larger discrepancies for some measurements (reaching deviations of 10 K) mainly during wintertime. Differences between TEMPERA and MLS retrievals can arise from several factors, including differences due to spatio-temporal inhomogeneities arising from synoptic variability which can be more important during winter, differences in

15 vertical resolution, interpolation techniques, or measurements errors from both instruments.

The mean and the standard deviations of the difference between the TEMPERA and MLS measurements for both periods described in the previous section and also for the different seasons have been plotted in Figure 10. From this comparison, a clear change in the mean bias is again observed between both periods in the lower part of the stratosphere (from 20 to 37 km). In that range, the mean bias in period 1 was 2.5 ± 1.3 K, reaching a maximum deviation of 4.1 K at 28.5 km, while for period

20 2 the mean bias was -0.4 ± 0.9 K with a maximum negative deviation of -1.4 K at 30 km. In the upper part (between 38 and 50 km) the differences in the biases were not so significant with a mean value of 1.7 ± 0.5 K for period 1 and 2.3 ± 0.7 K for period 2. The standard deviations again show higher values for period 1 than for period 2, although the differences were smaller than in the comparison with RS. The mean standard deviations in the range between 20 and 50 km were 2.4 ± 0.6 K for period 1 and 2.0 ± 0.4 K for period 2.

25 This comparison also shows a seasonal behaviour for the mean and the standard deviation of the temperature differences between TEMPERA and MLS for both periods. For period 1 there was a positive bias for both seasons in almost the whole column with larger values in winter than in summer. The mean bias in the lower part (20-35 km) was 3.3 ± 1.2 K in winter and 1.9 ± 1.4 K in summer. The discrepancies were even larger in the upper part (35-50 km) showing a much lower bias in summer (0.4 ± 0.4 K) than in winter (2.8 ± 0.7 K). During period 2 the differences between the biases in winter and summer were quite

30 constant in altitude and they were always lower than 1.6 K. The standard deviations of the temperature differences showed higher values in winter than in summer for both periods. For period 1 the mean standard deviation for the whole range (20-50 km) was 2.5 ± 0.5 K in winter reaching a maximum value (3.1 K) at 28.5 km while for period 2 the mean standard deviation was 2.1 ± 0.5 K with a maximum value of 2.6 K at 32 km. The standard deviations in summer for both periods were very similar with mean values for the whole altitude range (20-50 km) of 1.8 ± 0.6 K in period 1 and 1.7 ± 0.5 K in period 2. These results

35 again show the lower temperature discrepancies observed between TEMPERA and the MLS satellite during summertime. The
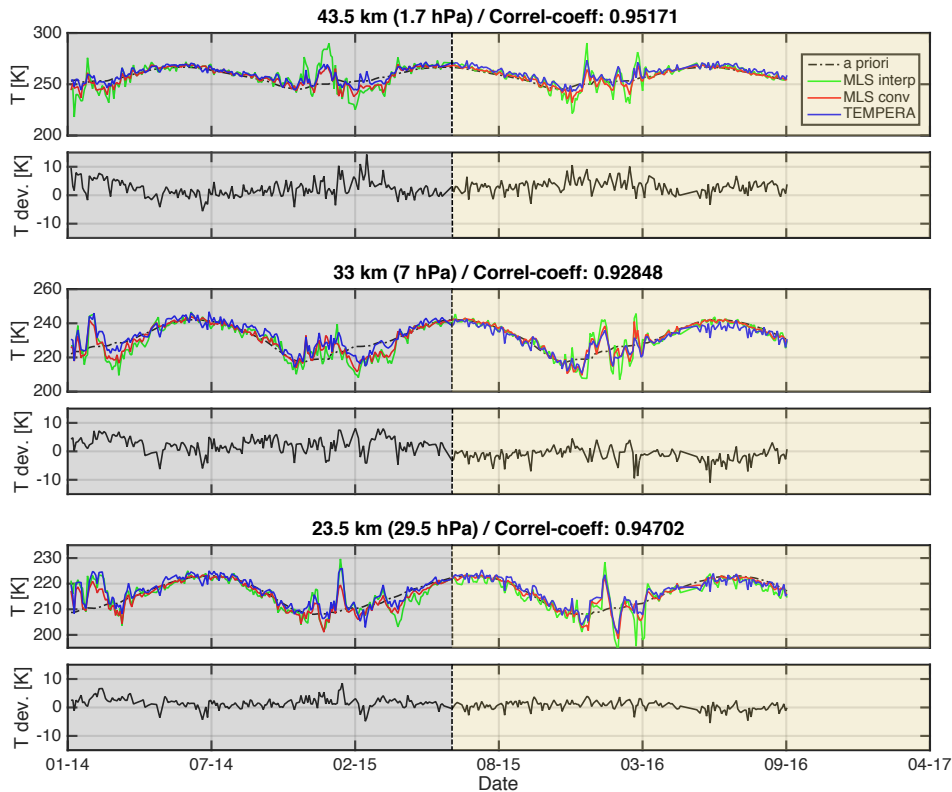
Figure 9. Stratospheric temperature evolution and temperature deviations at different altitudes for TEMPERA and MLS. Different background colors are used to distinguish between period 1 and 2 (gray and light brown, respectively).

biases found in this comparison are similar to those reported by Schwartz et al. (2008) for a comparison between MLS version 2.2 retrievals and different analyses and observations (GEOS-5, ECMWF, radiosondes, AIRS/AMSU, etc), where the biases ranged between -2.5 K and 1 K.

The MLS measurements have also been compared with the ones from RS in the range where they were comparable (lower stratosphere). Only collocated MLS profiles (according to the criteria as used above) and measured within 4 hours of the RS launch were selected for the comparison. A total of 323 pairs of profiles fulfilled these criteria and were used for these statistics. The RS profiles were interpolated to the pressure grid of MLS in order to perform the direct comparison of their profiles. Figure 11 shows the mean and the standard deviation for this comparison. We can observe that the mean bias ranges from -1.7 K at 19 km to +1.4 at 15 km. The standard deviation of the temperature differences between MLS and RS was quite constant with altitude, with a mean value of 1.7 ± 0.2 K and a maximum value of 2.2 K reached at 31 km. We note that the bias and the standard deviation observed between MLS and RS are very similar to the values observed in the comparison between TEMPERA and RS in period 2 (biases ranging between -2.3 K and 0.9 K and the standard deviations between 1.3 K and 1.7
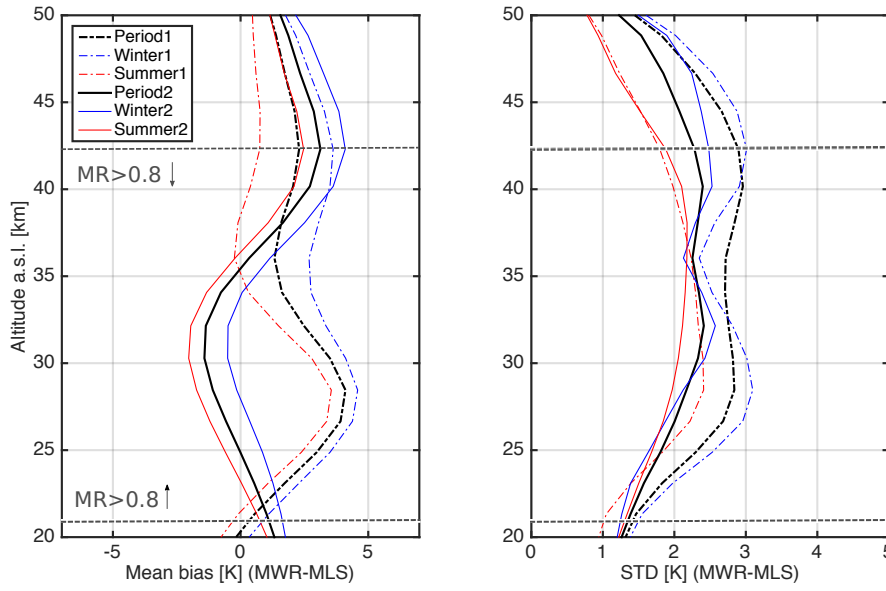
Figure 10. Mean temperature biases and standard deviations between TEMPERA and MLS. A total of 358 profiles have been compared (Period 1: 192 prof., dash lines; Period 2: 166 prof., solid lines). The mean and the standard deviations for each period are represented by black lines. The winter season is indicated with blue lines while the summer is indicated by red lines (Winter1: 103 prof.; Summer1: 89 prof.; Winter2: 67 prof. Summer2: 99).

K). The slight underestimation of the temperature in most of the altitudes found for MLS versus RS in this study agrees with the results obtained by Schwartz et al. (2008) between MLS and different sources.

4.3 Comparison with lidar measurements

The TEMPERA radiometer has also been compared with an active remote sensing instrument, a Rayleigh lidar. This lidar is
5 operated at Hohenpeißenberg station (Germany) which is located around 400 km northwest of Payerne. Despite the distance between both instruments, we wanted to evaluate the agreement in the stratospheric temperature between these very different techniques. A total of 192 profiles have been compared for all weather conditions (except for rainy cases) for the period from January 2014 to July 2016. As in the previous comparisons the lidar profiles were interpolated to the pressure grid of TEMPERA radiometer and then these profiles were convolved using the averaging kernel of TEMPERA. Since the Rayleigh
10 lidar only provides temperature information above approximately 28 km (below this the measurements would be affected by stratospheric aerosol), the gap below this altitude was filled with coincident measurements from TEMPERA in order to avoid modifying the averaging kernel used by TEMPERA for the convolution.

Figure 12 shows the stratospheric temperature evolution from TEMPERA and the lidar at three different altitude levels. For the lowest altitudes shown here (29.5 km, asl), the temperature from RS has also been plotted since at this altitude there were
15 measurements from the three instruments. We can observe from this figure that there is good agreement between TEMPERA
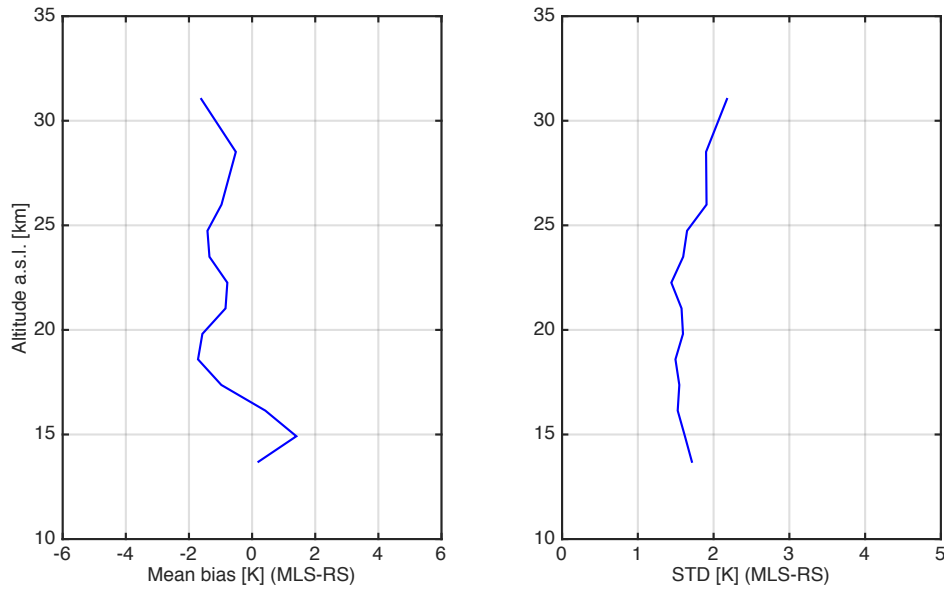
Figure 11. Mean bias and standard temperature deviation between MLS and RS.

and the lidar in the upper stratosphere, with correlation coefficients larger than 0.94 for the two highest altitudes. This coefficient is lower (0.9) for the lowest altitude (29.5 km, asl). The agreement between the lidar and the RS in this lowest altitude is better than for TEMPERA with a correlation coefficient of 0.96. The evolution of the temperature deviations between TEMPERA and lidar at the three altitudes shows small discrepancies for both techniques over the measurement period, with the values in most measurements below 5 K. The biggest differences were found at the lowest altitude (29.5 km, asl), where a clear change of bias was observed after summer 2015.

Figure 13 shows the mean bias and the standard deviation for all the measurements in periods 1 and 2 in addition to seasonal profiles. Mean bias profiles show again a clear change in the tendency of the biases of both periods, being more evident in the lower stratosphere (below 35 km). In this lowest altitude range the mean biases were 2.7±1.3 K for period 1 and -1.2±0.4 K for period 2. Above 35 km the differences between the biases were smaller with a larger bias for period 2 (2.3±0.9 K versus 1.3±0.4 K in period 1). Similar behaviour to the other comparisons has been observed for the standard deviation with larger values during period 1 than during period 2. The mean values for the whole altitude range were 2.9±0.3 K for period 1 and 2.5±0.2 K for period 2. Seasonal behaviour is observed in the bias and standard deviation for both periods. The seasonal biases showed a vertical oscillation with different tendencies for both periods in the lower and upper part of the stratosphere. For the lower part (28-35 km) the mean biases for period 1 (period 2) were 3.2±1.1 K (-0.7±0.4 K) in winter and 1.9±1.5 K (-2.1±0.3 K) in summer. In the upper part (35-50 km), a general positive bias was observed between TEMPERA and the lidar where the mean biases for period 1 (period 2) were 2.2±0.6 K (2.9±1.1 K) in winter and -0.3±0.3 K (1.1±0.9 K) in summer. The standard deviations showed larger values in winter for both periods than in summer. The highest standard deviations were again
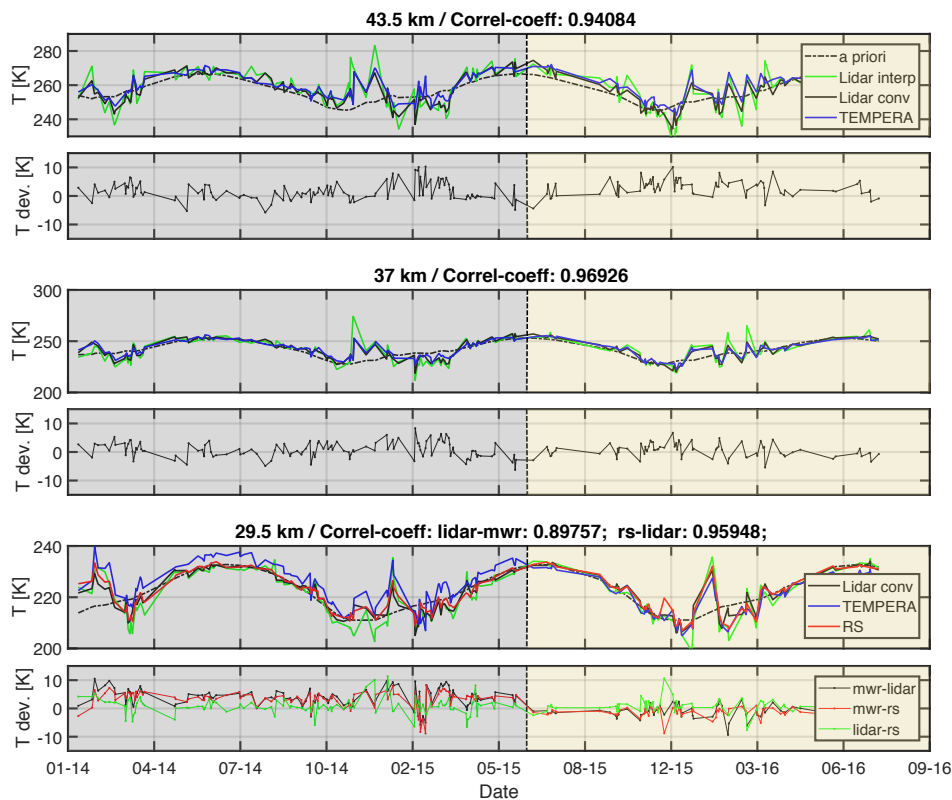
Figure 12. Stratospheric temperature evolution from TEMPERA, lidar and RS. Different background colors are used to distinguish between period 1 and 2 (gray and light brown, respectively).

observed in the winter of period 1. The mean standard deviations in the whole column for period 1 (period 2) were 3.1±0.4 K (2.6±0.3 K) in winter and 2.0±0.3 K (1.7±0.4 K) in summer.

4.4 Comparison with SD-WACCM

A first validation of the stratospheric temperature from SD-WACCM (Whole Atmosphere Community Climate Model with
5 Specified Dynamics) has also been carried out in this study. SD-WACCM is the whole atmosphere component of CESM (Community Earth System Model) (Kunz et al., 2011; Lamarque et al., 2012). CESM is a coupled climate model which means that it consists of separate models for different parts of the climate system which interact via the coupler module. There are models for ocean, atmosphere, land, sea ice, land ice and rivers. CESM allows to combine the above models to a component set for the simulation.
10 The Specified Dynamics (SD) used in these simulations means that the model is nudged by meteorological analysis fields by 10% at every internal time-step up to an altitude of 50 km. This means 90% of the model and 10% of the nudging data are
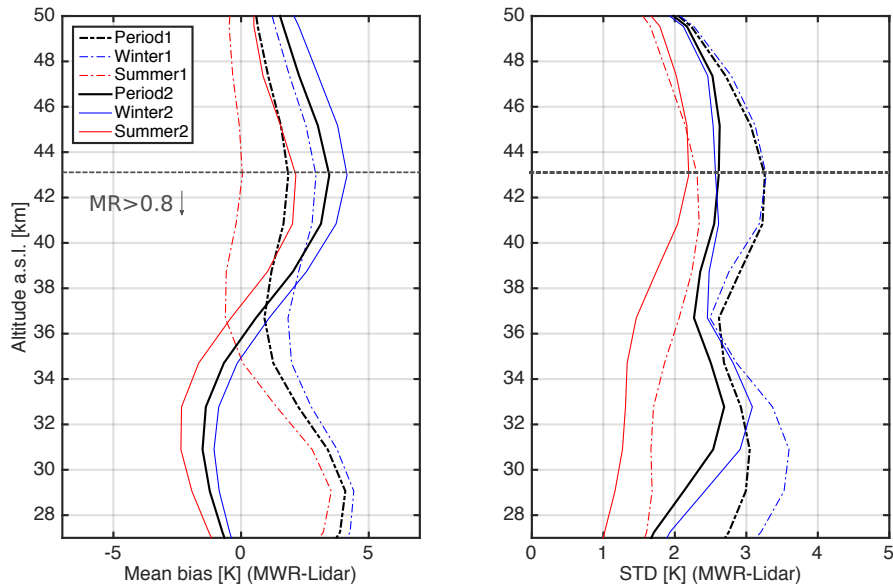
Figure 13. Mean temperature deviation between TEMPERA and lidar. A total of 192 profiles have been compared (Period 1: 117 prof., dash lines; Period 2: 75 prof., solid lines). The mean bias and the standard deviations for each period are represented by black lines. The winter season is indicated with blue lines while the summer is indicated by red lines (Winter1: 73 prof.; Summer1: 44 prof.; Winter2: 49 prof. Summer2: 26).

taken. The fields that are nudged are temperature, horizontal winds, surface wind stress, surface pressure and heat fluxes from the surface. The nudging data are from the Goddard Earth Observing System version 5.0.1 (GEOS-5) Data Assimilation and is provided every 6 hours, in between the data are interpolated.

The altitude range for SD-WACCM is from ground to 140 km (asl). The altitude resolution ranges from 0.5 to 4 km (with
5 lower resolution at higher levels) and with a total of 88 layers in the whole atmosphere. The grid resolution is 1.9° latitude by 2.5° longitude.

The stratospheric temperatures from SD-WACCM have been compared with the almost continuous stratospheric temperature profiles measured by the TEMPERA radiometer for the period from January 2014 to April 2016. A total of 6868 profiles were selected for comparison under all weather conditions except rainy conditions. Figure 14 shows the stratospheric temperature
10 evolution along this period for TEMPERA and WACCM. Good agreement is observed in general between both temperature sets. We can observe that the temperature from the model follows the same pattern as TEMPERA, with the same annual cycle and detecting the same structures in time and also in altitude. Note the good agreement observed during winters, where strong increases in temperatures are produced for short periods and can be observed in both data sets. The differences between TEMPERA and WACCM are more evident above 50 km (asl), but above this altitude the measurement response for TEMPERA
15 is low (lower than 0.6) since the weight of the measurements is small and so it should not be considered in the comparison.
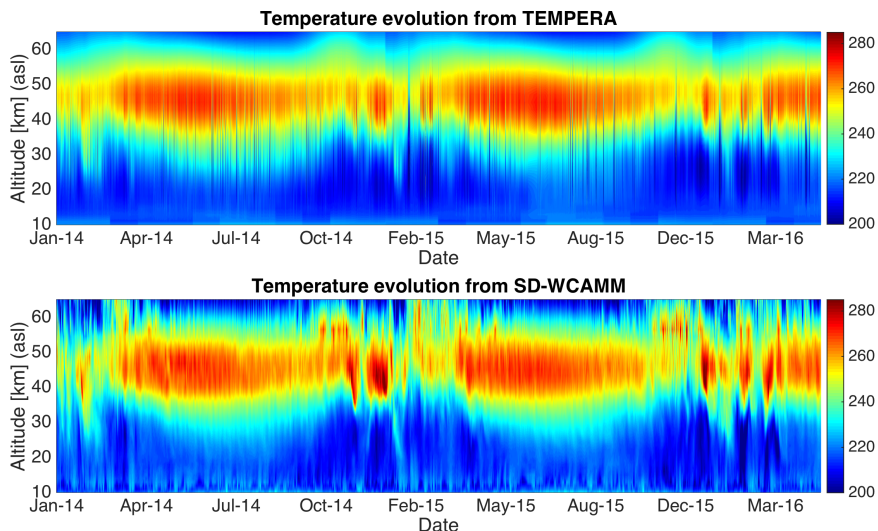
Figure 14. Stratospheric temperature from TEMPERA radiometer (upper panel) and WACCM model (lower panel).

The temperature profiles from SD-WACCM have been interpolated and convolved as described in section 3 to allow comparison with those from TEMPERA. Figure 15 shows the evolution of the temperature at three altitude levels and the differences between both (TEMPERA-WACCM). The good agreement observed from these plots is particularly shown by the low temperature deviation values (lower than 5 K most of the time) and the large correlation coefficient (larger than 0.92). Despite
5 this good agreement, we also find some periods with larger discrepancies between the measurements and the model, especially during winter time, most markedly in winter 2015. Note that the statistics shown in this section, are particularly robust, since almost 7000 pairs of temperature profiles are compared.

We have also calculated the bias and the standard deviation for this comparison between the TEMPERA radiometer and the WACCM model (Fig. 16). It is again very obvious from these statistics that there is a strong change in the biases between
10 periods 1 and 2, with a very different tendency in the lower stratosphere than in the upper. The mean biases for the lower part (20-35 km) were 1.4±1.1 K for period 1 and -1.0±1.3 K for period 2, whilst the mean biases for the upper stratosphere (35-50 km) were 1.0±0.7 K for period 1 and 1.7±1.1 K for period 2. The seasonal behaviour observed in the biases was almost negligible for both periods.

From the standard deviation figure (Fig. 16, right) we can observe that much larger values are obtained for period 1, with
15 a mean value in the whole column of 2.9±0.6 K and a maximum standard deviation of 3.8 K at 29 km. These large standard deviations observed during period 1 are strongly influenced by the large values observed during winter time (blue dashed line), when a maximum mean standard deviation of 4.7 K at 29 km was reached. The rest of the standard deviation profiles show very similar values between them, increasing slightly in the lower part (up to 30 km), and keeping close to constant values above this altitude. The smallest values are found in summer with a mean bias in the whole column of 1.8±0.4 K for period 1
20 and 1.5±0.3 K for period 2.
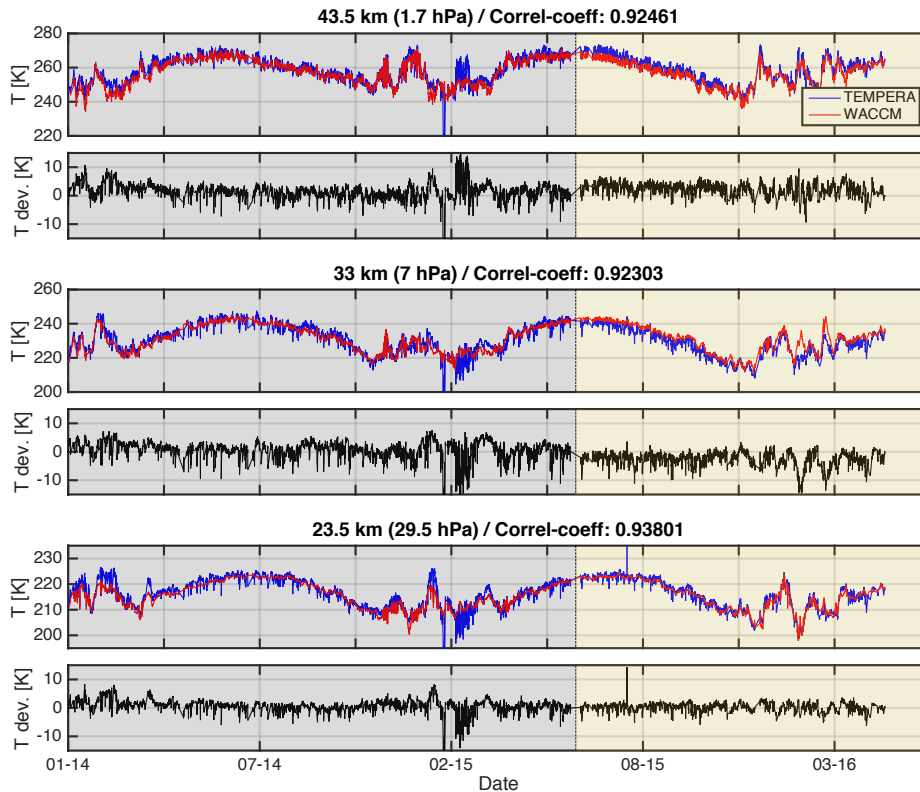
Figure 15. Stratospheric temperature evolution from TEMPERA and WACCM. Different background colors are used to distinguish between period 1 and 2 (gray and light brown, respectively).

4.5 All measurements and model versus TEMPERA

In order to summarize the intercomparison carried out between TEMPERA and the different measurement techniques and model we have plotted the biases and the standard deviations for all the comparisons together (Fig. 17). Since we are interested in evaluating the accuracy and precision of TEMPERA radiometer against other measurements in this study we have only dis-
5 played in Fig. 17 the biases and the standard deviations obtained for the summer season, since it is less affected by atmospheric variability than the winter measurements.

The mean bias plot (Fig. 17, left) shows a clear change of biases between TEMPERA and all the other measurements between the first (dashed lines) and the second (solid lines) period (before and after the repair of the FFT spectrometer's attenuator). We can observe that there is a persistent vertical oscillation for all the profiles in both periods causing a different behaviour of
10 the biases in the lower and upper stratosphere. This oscillation has an amplitude of around 2 K and a periodicity of roughly 20 km. Similar behaviour was observed for the MLS measurements when they were compared with different sources (Schwartz et al., 2008). The change of tendency in the bias between both periods is more evident in the lower stratosphere (below 35 km) where we can observe that for almost all the altitude levels the biases change from positive to negative in all the comparisons.
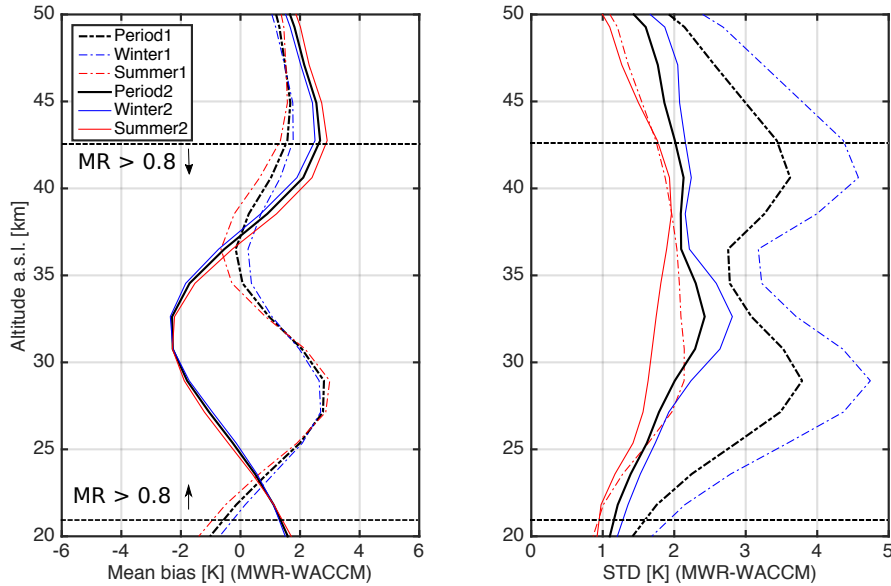
Figure 16. Mean temperature deviation between TEMPERA and the WACCM model. A total of 6868 profiles have been compared (Period 1: 4339 prof., dash lines; Period 2: 2529 prof., solid lines). The mean and the standard deviations for each period are represented by black lines. The winter season is indicated by blue lines while the summer is indicated by red lines (Winter1: 2361 prof.; Summer1: 1978 prof.; Winter2: 1473 prof. Summer2: 1056).

Another remarkable point is the consistency between the different biases in each period, showing small differences between them (below 1K) for most of the altitudes, especially for period 2. For period 1, the maximum deviation was found at 28.5 km, with a maximum value of 3.6 K for the comparison with the MLS satellite. Below this altitude, an almost identical bias between the comparison with RS and WACCM model is found. In the upper stratosphere the biases were between -0.6 K and 5 1.5 K, and the smallest bias was found in the lidar comparison. For period 2 the values of the different biases ranged between -2.4 K (at 32 km) and a maximum positive bias of 2.9 K (at 43 km), the latter being found with the comparison with WACCM. As we already mention the differences between the different comparisons for period 2 were smaller than for period 1 showing consistency between the RS, MLS, lidar measurements and also WACCM simulations.

Fig. 17(right) shows the standard deviations of the differences between TEMPERA and the different measurements and 10 model. In general we can observe that there was a reduction in the standard deviations for all the comparisons in period 2, indicating that the precision of TEMPERA improved after the attenuator was repaired. Next, we will focus our discussion on period 2, when we consider that TEMPERA was operating optimally. For this period, we can observe that standard deviations were always lower than 2.2 K, with this maximum value being reached at 45 km in the comparison with the lidar. The lowest standard deviation in the lower stratosphere (< 35 km) was found for the comparison with RS, with the mean value in this 15 range being 1.3±0.1 K. The highest standard deviations in the lower stratosphere were found in the comparison with MLS (1.3±0.1 K). These results evidence a better precision from the TEMPERA radiometer when it is compared with the in situ

Table 1. Range of biases and standard deviations of the TEMPERA radiometer when compared with RS, MLS, lidar and WACCM measurements/model results.

		MWR-RS	MWR-MLS	MWR-lidar	MWR-WACCM
lower strat.	BIAS	-1.3±1.1	-1.0±1.0	-1.1±1.3	-1.0±1.3
(20-35 km)	STD	1.3±0.1	1.8±0.3	1.1±0.2	1.5±0.3
upper strato.	BIAS		1.5±0.9	1.1±0.9	1.9±1.1
(35-50 km)	STD		1.7±0.5	1.9±0.3	1.6±0.3

reference technique of RS in the lower stratosphere. This result makes sense since RS is the technique with the lowest errors (0.2 K for temperature) and the comparison between TEMPERA and RS is the one that should present lower atmospheric variability between both measurements since the RS are launched from the same location where TEMPERA is operated.

In the middle stratosphere (between 30 and 40 km) the lowest standard deviations were found for the comparison with lidar with a mean value of 1.4±0.2 K. However, above this altitude (40 km) the standard deviation with respect to lidar is the largest (2.2±0.3 K). In this upper part the lowest standard deviations were found for the comparison with MLS and WACCM. A common pattern that is observed in all the comparisons is that the standard deviations decrease slightly with altitude in the last kilometres of the stratosphere. This behaviour is due to a greater weight of the a priori temperature profile used in the TEMPERA retrievals and also in the convolved profiles in these altitudes since the measurement response presents lower values for high altitudes (around 0.6 at 48 km).

Table 1 presents the different biases and standard deviations obtained in the lower and upper stratosphere for all the comparisons during summertime in period 2. These values are the most representative way of characterizing the accuracy and precision of the TEMPERA radiometer since they correspond to the period when TEMPERA was running with the repaired attenuator (period 2) and also when the measurements where least affected by atmospheric variability (summertime).

We end by highlighting the consistency found between the standard deviations of the different comparisons and the observation errors of the TEMPERA retrievals. As mentioned in section 3.1 the OEM also estimates the observation, smoothing and total errors of the TEMPERA inversions (Fig. 4e). The standard deviations found in the different comparisons are partly related to the observation error of TEMPERA but also to the errors associated with the other measurements and the atmospheric mismatches. If we assume that the random errors in TEMPERA (D1), in the other instruments (D2) and in the atmospheric mismatching (D3) are independent, then the observed standard deviation (DT) should be given by $DT^2 = DT1^2 + DT2^2 + DT3^2$. For example, if we consider the observation error of TEMPERA provided by OEM (0.8 K), the errors for the lidar (0.7 K) and the mean observed standard deviation for the comparison between TEMPERA and the lidar (1.1 K) we would conclude that the errors associated with atmospheric mismatches should be 0.3 K, which is a realistic value and shows the consistency between the observed standard deviations and the observation errors of the different measurements.
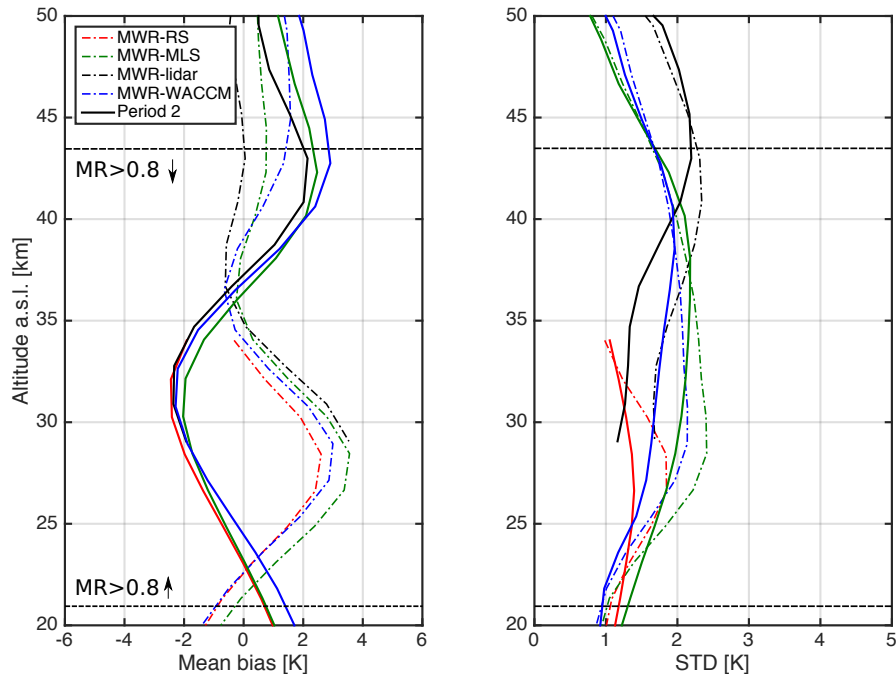
Figure 17. Mean and standard temperature deviations between the TEMPERA radiometer and the measurements from the different instruments and WACCM.

5 Conclusions

Nearly three years of measurements of stratospheric temperature profiles from a relatively new ground-based microwave radiometer (TEMPERA) have been intercompared with those from different measurement techniques: RS, MLS satellite and Rayleigh lidar and also from the SD-WACCM model. TEMPERA measurements were carried out at the aerological station

5 of MeteoSwiss in Payerne from January 2014 to September 2016. Ground-based microwave measurements offer the advantages that they can provide unattended continuous measurements of temperature profiles in almost all weather conditions with a reasonably good spatial and temporal resolution. The stratospheric temperature profiles (from 20 to 50 km) were obtained from TEMPERA measurements using OEM by means of the radiative transfer model ARTS/QPack. All the profiles from the other techniques (RS, MLS and lidar) and from the WACCM model were interpolated to the TEMPERA pressure grid and then

10 convolved using the averaging kernel of this radiometer in order to be compared with the profiles from TEMPERA.

The temperature evolutions measured at different altitudes by TEMPERA and the other techniques, as well as the model, showed in general a very good agreement with a high correlation (always larger than 0.9) between the datasets. The stratospheric temperature evolutions showed a larger variability during wintertime and also evidenced larger discrepancies between TEMPERA and the other datasets during those periods. A small step in the temperature deviations was observed in July of

15 2015 for the different comparisons, which was related to the repair of an attenuator in the FFT spectrometer of TEMPERA.

This repair caused a small change in the measured brightness temperature from TEMPERA and therefore in the retrieved temperature profile. For this reason, and in order to take into account the instrument modification and characterize possible changes in the accuracy and precision of TEMPERA radiometer, the statistical analysis was carried out over two different measurement periods (before and after the modification). In addition a seasonal distinction (winter and summer) was considered in the statistics to take into account the larger atmospheric variability that can be observed during wintertime and which could produce larger deviations between the instruments due to the atmospheric conditions.

The accuracy and the precision of the TEMPERA radiometer has been evaluated by means of the bias relative to other measurement techniques and model output (RS, MLS, lidar and WACCM), as well as the standard deviation of the difference between the measurements. The stratospheric temperature comparison between TEMPERA and the other datasets showed a clear change in the biases between periods 1 and 2 (before and after the repair of the attenuator) in all the statistics. For the lower stratosphere (20-35 km) the biases changed from positive values in period 1 to negative values in period 2. The smallest mean deviations were observed in the comparison with RS, with values always lower than ± 2.5 K. The largest biases were observed for the comparisons with MLS and the Rayleigh lidar reaching maximum deviations of around +4.5 K at some altitudes in period 2. In general the biases were smaller and negative for all the comparisons during period 2, indicating a slight underestimation of the temperature by TEMPERA radiometer in that period.

In the upper part of the stratosphere (above 35 km) the differences between both periods were not so evident, and generally positive biases were observed in both periods for all the comparisons. The deviations in this upper part were always less than 4.5 K. We would like to point out that only weak seasonal behaviour was observed for the biases in the comparisons with RS and WACCM, whereas it was more pronounced for the comparison with MLS and the lidar, especially in period 1.

The standard deviations obtained from the different statistics showed very different results in the two periods. Larger values were observed for all the comparisons in period 1 than in period 2, indicating that the precision of TEMPERA radiometer improved after the repair of the spectrometer's attenuator. The standard deviations were especially high in wintertime during period 1, reaching maximum values of around 4.5 K for the comparison with RS (at 28 km) and MLS (28 km and 41 km). In period 2 the standard deviations during winter were also larger than in summer but with smaller differences (except for the lidar in the lower part of the stratosphere). These results confirmed the larger atmospheric variability that can be found during wintertime and which produces a lower agreement in the temperature measurements between the different instruments, especially when the horizontal distance between them is large.

Finally, the accuracy and the precision of the TEMPERA radiometer have been characterized by means of the bias relative to the different measurement and model values, as well as by the standard deviations of temperature differences between TEMPERA and the other values. All of this was done during period 2 (instrument in optimal conditions) and in summer (less affected by atmospheric variability). These statistics in the lower stratosphere (below 35 km) showed mean biases ranging between 1.0 and 1.3 K (max. for RS and min. for MLS) and mean standard deviations that ranged between 1.1 and 1.8 K (max. for MLS and min. for lidar). While in the upper stratosphere (above 35 km) the mean biases ranged between 1.1 and 1.9 K (max. for WACCM and min. for lidar) and the mean standard deviations ranged between 1.6 and 1.9 K (min. for WACCM and

max. for lidar). The standard deviations observed in the different comparisons were consistent with the observation errors that are expected from the different instruments, indicating that it is a good measure of the instrument errors.

From all these intercomparisons we can conclude that the TEMPERA radiometer performed well at determining temperatures in the stratosphere. Continuous TEMPERA measurements will in the future make it possible to carry out temperature trend analyses, which are an important component of global change. These trends can provide evidence of the roles of natural and anthropogenic climate change mechanisms. Stratospheric temperature changes are also crucial for understanding stratospheric ozone variability and trends, including predicting future changes. In addition, measurements with a high temporal resolution in a fixed location will make it possible to characterize the local thermodynamics, which can be especially interesting during wintertime.

Acknowledgements. We thank MeteoSwiss and in particular Dominique Ruffieux, Ludovic Renaud, Philippe Overney and Jean-Marc Aellen for hosting our instrument and for the support on-site. We would also like to thank to Dr. Peter Speirs for his contribution with the language revision of this manuscript. This work has been funded by the Swiss National Science Foundation under grant 200020-160048 and MeteoSwiss in the framework of the GAW project "Fundamental GAW Parameters by Microwave Radiometry".

References

Anderson, G. P., Clough, S., Kneizys, F., Chetwynd, J., and Shettle, E. P.: AFGL atmospheric constituent profiles (0.120 km), Tech. rep., DTIC Document, 1986.

Aumann, H. H., Chahine, M. T., Gautier, C., Goldberg, M. D., Kalnay, E., McMillin, L. M., Revercomb, H., Rosenkranz, P. W., Smith, W. L., Staelin, D. H., et al.: AIRS/AMSU/HSB on the Aqua mission: Design, science objectives, data products, and processing systems, IEEE Transactions on Geoscience and Remote Sensing, 41, 253–264, 2003.

Bindoff, N. L., Stott, P. A., AchutaRao, M., Allen, M. R., Gillett, N., Gutzler, D., Hansingo, K., Hegerl, G., Hu, Y., Jain, S., et al.: Detection and attribution of climate change: from global to regional, 2013.

Bleisch, R., Kämpfer, N., and Haefele, A.: Retrieval of tropospheric water vapour by using spectra of a 22 GHz radiometer, Atmospheric Measurement Techniques, 4, 1891–1903, doi:10.5194/amt-4-1891-2011, 2011.

Eriksson, P., Buehler, S., Davis, C., Emde, C., and Lemke, O.: ARTS, the atmospheric radiative transfer simulator, version 2, Journal of Quantitative Spectroscopy and Radiative Transfer, 112, 1551–1558, 2011.

Flury, T., Hocke, K., Haefele, A., Kämpfer, N., and Lehmann, R.: Ozone depletion, water vapor increase, and PSC generation at midlatitudes by the 2008 major stratospheric warming, Journal of Geophysical Research: Atmospheres, 114, 2009.

Haefele, A., De Wachter, E., Hocke, K., Kämpfer, N., Nedoluha, G., Gomez, R., Eriksson, P., Forkman, P., Lambert, A., and Schwartz, M.: Validation of ground-based microwave radiometers at 22 GHz for stratospheric and mesospheric water vapor, Journal of Geophysical Research: Atmospheres, 114, 2009.

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

Keckhut, P., Wild, J. D., Gelman, M., Miller, A. J., and Hauchecorne, A.: Investigations on long-term temperature changes in the upper stratosphere using lidar data and NCEP analyses, Journal of Geophysical Research: Atmospheres, 106, 7937–7944, 2001.

Kunz, A., Pan, L., Konopka, P., Kinnison, D., and Tilmes, S.: Chemical and dynamical discontinuity at the extratropical tropopause based on START08 and WACCM analyses, Journal of Geophysical Research: Atmospheres, 116, 2011.

Lamarque, J.-F., Emmons, L., Hess, P., Kinnison, D. E., Tilmes, S., Vitt, F., Heald, C., Holland, E. A., Lauritzen, P., Neu, J., et al.: CAM-chem: Description and evaluation of interactive atmospheric chemistry in the Community Earth System Model, Geoscientific Model Development, 5, 369, 2012.

Liebe, H., Hufford, G., and Cotton, M.: Propagation modeling of moist air and suspended water/ice particles at frequencies below 1000 GHz, in: In AGARD, Atmospheric Propagation Effects Through Natural and Man-Made Obscurants for Visible to MM-Wave Radiation 11 p (SEE N94-30495 08-32), vol. 1, 1993.

Moreira, L., Hocke, K., Eckert, E., von Clarmann, T., and Kämpfer, N.: Trend analysis of the 20-year time series of stratospheric ozone profiles observed by the GROMOS microwave radiometer at Bern, Atmospheric Chemistry and Physics, 15, 10 999–11 009, 2015.

Navas-Guzmán, F., Stähli, O., and Kämpfer, N.: An integrated approach toward the incorporation of clouds in the temperature retrievals from microwave measurements, Atmospheric Measurement Techniques, 7, 1619–1628, 2014.

Navas-Guzmán, F., Kämpfer, N., Murk, A., Larsson, R., Buehler, S., and Eriksson, P.: Zeeman effect in atmospheric O2 measured by ground-based microwave radiometry, Atmospheric Measurement Techniques (AMT), 8, 1863–1874, 2015.

Navas-Guzmán, F., Kämpfer, N., and Haefele, A.: Validation of brightness and physical temperature from two scanning microwave radiometers in the 60 GHz O2-band using radiosonde measurements, Atmospheric Measurement Techniques, 9, 4587–4600, 2016.

Ramaswamy, V. and Schwarzkopf, M.: Effects of ozone and well-mixed gases on annual-mean stratospheric temperature trends, Geophysical research letters, 29, 2002.

Randel, W. J., Shine, K. P., Austin, J., Barnett, J., Claud, C., Gillett, N. P., Keckhut, P., Langematz, U., Lin, R., Long, C., et al.: An update of observed stratospheric temperature trends, Journal of Geophysical Research: Atmospheres, 114, 2009.

5 Remsberg, E., Lingenfelser, G., Harvey, V., Grose, W., Russell, J., Mlynczak, M., Gordley, L., and Marshall, B.: On the verification of the quality of SABER temperature, geopotential height, and wind fields by comparison with Met Office assimilated analyses, Journal of Geophysical Research: Atmospheres, 108, 2003.

Rodgers, C. D.: Inverse methods for atmospheric sounding: theory and practice, vol. 2, World scientific, 2000.

Rosenkranz, P.: Absorption of microwaves by atmospheric gases, 1993.

10 Rosenkranz, P. W.: Water vapor microwave continuum absorption: A comparison of measurements and models, Radio Science, 33, 919–928, 1998.

Santer, B., Penner, J., Thorne, P., Collins, W., Dixon, K., Delworth, T., Doutriaux, C., Folland, C., Forest, C., Hansen, J., et al.: How Well Can the Observed Vertical Temperature Changes be Reconciled with our Understanding of the Causes of These Temperature Changes?, 2006.

15 Scheiben, D., Straub, C., Hocke, K., Forkman, P., and Kämpfer, N.: Observations of middle atmospheric H_2O and O_3 during the 2010 major sudden stratospheric warming by a network of microwave radiometers, Atmospheric Chemistry and Physics, 12, 7753–7765, 2012.

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

20 Schwarzkopf, M. D. and Ramaswamy, V.: Evolution of stratospheric temperature in the 20th century, Geophysical Research Letters, 35, 2008.

Shvetsov, A., Fedoseev, L., Karashtin, D., Bol'shakov, O., Mukhin, D., Skalyga, N., and Feigin, A.: Measurement of the middle-atmosphere temperature profile using a ground-based spectroradiometer facility, Radiophysics and Quantum Electronics, 53, 321–325, 2010.

Stähli, O., Murk, A., Kämpfer, N., Mätzler, C., and Eriksson, P.: Microwave radiometer to retrieve temperature profiles from the surface to 25 the stratopause, Atmospheric Measurement Techniques, 6, 2477–2494, 2013.

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

Waters, J. W.: Ground-based measurement of millimetre-wavelength emission by upper stratospheric O_2, Nature, 242, 506–508, 1973.

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

Yan, X., Wright, J. S., Zheng, X., Livesey, N. J., Vömel, H., and Zhou, X.: Validation of Aura MLS retrievals of temperature, water vapour and ozone in the upper troposphere and lower-middle stratosphere over the Tibetan Plateau during boreal summer, Atmospheric Measurement 35 Techniques, 9, 3547, 2016.