AMT-2019-455

Measurement Characteristics of an airborne Microwave Temperature Profiler (MTP)

by Mareike Heckl et al.

Reply to the referee #1 comments

We would like to thank the referee for again thoroughly reviewing our manuscript and for the helpful advice provided in the comments, below, which we believe helped further improving the manuscript. In the following we give our answers regarding the points made by the reviewer. The statements, comments and suggested corrections raised by the referee are printed in black italics and our comments are presented in blue. We tried to consider all of the raised points in the revised manuscript in an adequate manner.

Answers to referee #1

The manuscript has improved significantly, however, there are some flaws that need to be corrected. As a major issue a clear definition and subsequent wording of the terms how the measurement uncertainty characterized in terms of systematic and random components needs to be made (see also comments below). In this respect it might also be important to mention the uncertainty characteristics of the in-situ HALO-TS measurement which is given as 0.5 K. However, the standard deviation between HALO-TS and TB is ≤ 0.38 K which implies – together with the random noise characteristics of TB (around 0.3 K) that the precision (noise) of HALO-TS is much lower than that.

We agree that the revised manuscript has improved. In the newly revised version, we have re-worded the text where necessary, according to the comments below, and making sure that the definition of noise characterisation is understandable and clear throughout the manuscript.
Concerning the noise characterisation of HALO-TS, the authors have cited available literature, which states the overall-error of HALO-TS (including accuracy). We support the observation that the precision of HALO-TS is better than that, (also implied by the overall accuracy of the sensors used to measure HALO-TS; see Ungermann et al., 2015, and supported by the observed RMS difference to the mean value of the 13s interval used in our study; see Table 7), however, we want to refrain from stating this in our manuscript without further investigation, and keep the cited value of 0.5 K overall uncertainty.

The term "noise figure" is used frequently in the text at several instances when a general characterization of the noise is meant. However, the term "noise figure" is a typical specification for microwave elements (mixers, amplifiers) expressing the noise factor between in and output in decibels (given in dB). Therefore, please use this term with care. Most of the time it will be sufficient to just remove the word "figure" or replace it with "characterisation".

We have followed the reviewer's suggestion to remove the word "figure" in most cases, and replacing it with "characterisation" in some instances. The changes can be seen in the marked-up document.

Line 22. Another important application of HAMP is the liquid water path (see Jacob et al., AMT, 2020). In fact later on it would be good to mention that also non-resonant interaction of microwave radiation with atmospheric hydrometeors influences TB.

The ability of HAMP to measure the LWP was added in the sentence with a reference to the paper by Jacob et al., 2019: https://doi.org/10.5194/amt-12-3237-2019.
Concerning the interaction with hydrometeors: We acknowledge the fact that hydrometeors can influence the measured TB, for example in case of aircraft icing. However, such situations are constantly avoided in flight. We also want to remind the referee that HALO is a high-altitude aircraft, with flight altitudes mostly above the tropopause, where water vapour concentrations are low. Still, the influence of hydrometeors is certainly important in the set-up of the retrieval algorithm, and should be discussed in that regard.

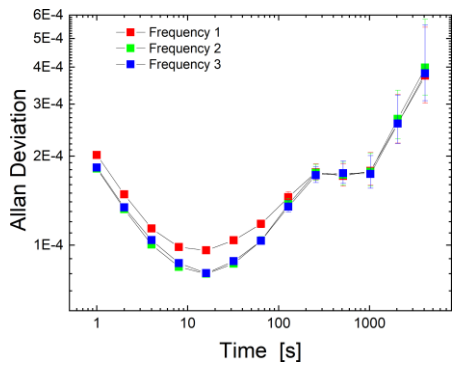
P2,l2: Shorten to "Based on these mesoscale temperature fluctuation analyses, a number of modelling studies aimed at improving the understanding and numerical description of atmospheric gravity waves"

We have made the suggested change.

P3l15: I would always been careful to make such a general statement as:
"For the first time, this study presents all relevant instrument characteristics of the HALO-MTP instrument." Be aware that there might be always surprises ahead of the road. For example, I think that there is currently no chance to specify the absolute accuracy of the MTP as there is no absolute truth available – HALO TS is only an approximation as the authors discuss later on. Furthermore, in terms of spectral analysis investigations the current investigation is not all-encompassing. The Allan variance could still provide additional information especially in terms of the maximum time between two subsequent calibrations.

The sentence was changed to: "For the first time, this study presents a thorough investigation of relevant instrument characteristics of the HALO-MTP instrument". We appreciate the fact that advances in science might lead to the insight that further characteristics have to be tested in order to fully understand MTP measurements – insights that might not be known today.
We do expect that the given uncertainty of MTP measurements provides a solid basis for confident and correct interpretation of MTP measurements (future and existing time series).

Concerning the Allen variance, we have pointed out, during the first review phase, that calibration measurements are already made as frequently as possible; i.e. during every single measurement cycle, corresponding to one calibration measurement every 13-14s, using standard measurement settings. As discussed previously, there is good reason to assume that this is already the optimum of calibration measurements given the necessary trade-off between the time it takes to record a measurement cycle (i.e. best-possible horizontal resolution of measurements) and best-possible knowledge of the instrument state. A quick-look of an Allen Variance analysis from the laboratory measurements indicates that the current settings fall within the ideal timing of calibration measurements.



P4l12: Please correct to "thermal radiation mainly emitted by oxygen .." Though the impact is low some influence of N2 and water vapor continuum occurrs especially when flying low. Similarly, hydrometeors can contribute to the signal.

We agree with the reviewer's comment, and have made the suggested change. We would like to mention that basic constituents mentioned by reviewer one are included in the radiative transfer calculations presented in the appendix.
A few lines later we have extended the sentence: "The recorded TBs have to be converted to absolute temperature profiles by using a retrieval algorithm that utilises forward radiative transfer calculations. " to: "The recorded TBs have to be converted to absolute temperature profiles by using a retrieval algorithm that utilises forward radiative transfer calculations, *in which all possible impacts on the measured radiance (e.g. emission by other trace gases such as water vapour or nitrous oxide, or hydrometeors,…) have to be considered.*"

P4, Eq. 2.1: Please mention that in this case the measured radiance I (usual notation) is equal to the Brightness B of a blackbody described by Plank's law. Also note that the omission of the higher order terms in the Rayleigh-Jeans approximation can be more than 0.5 K at your frequencies, check table 1 in Liu et al., 2008 https://doi.org/10.1016/j.jqsrt.2008.03.001

The information was added in the manuscript text, directly after the equation. The assumption that the measured signal has a linear relationship to the source temperature has been tested in the lab and it was shown that a calibration based on this assumption leads to reliable results in determining the brightness temperature. The point raised about using the Rayleigh-Jeans approximation rather than Planck's law again relates to radiative transfer calculations, which are part of the retrieval. Hence, the uncertainty in the Rayleigh-Jeans approximation and the corresponding error in the converted brightness temperature will not be explicitly discussed in this study.

P4, l26: The horn is no receiver. Please make also the connection to the antenna response function: "..a horn antenna guides the incoming atmospheric radiation and determines the spatial response function". What about the rotating mirror – is it just flat or does it focus the antenna beam?

We have corrected the sentence in the manuscript following the reviewer's suggestion, adding a few words concerning the mirror: "a horn antenna guides the incoming atmospheric radiation and, together with the hyperbolic shape of the rotating mirror, determines the spatial response function": The mirror is not flat, but is used to focus the antenna beam. On page 8 of the manuscript we already note that the antenna function is mainly defined by the shape of the rotating mirror. Also, in Table 1 we already point out that the mirror was explicitly designed to provide a beam width of 7.5 degrees of the antenna function. Here, it was now added, that the mirror has a hyperbolic shape.

P4, l28. Why don't you use the common term interim frequency (IF) instead of base-band? Further you could also provide the information that this is around 100 MHz which makes it much clearer for the reader in the following. Suggest the following text:
".. the incoming signal is converted to the interim frequency (IF) around 100 MHz. Both difference frequencies below and above the LO frequency are down-converted to the IF in the double side band receiver. Low pass filtering supresses any incoming radiation outside the IF bandwidth of 200 MHz such that the symmetric spectrum around the current LO frequency is measured with only a minor gap of approx. 20 MHz at the LO frequency. The IF signal is converted to a voltage.." In this way the information on Page 5 last line is not necessary. I think this information fits much better here – or is the IF bandwidth different for the wing-canister instrument than for other MTP?

We have made the suggested changes with the according information.
The HALO-MTP does indeed have slightly different operational settings than older versions of the MTP, which would be useful for the reader to know. The HALO-MTP does indeed operate at base-band, while older versions operated at 320 MHz (Mahoney, Denning, and Lim et al., priv. comm.).

P5, l2: "The PHYSICAL temperatures.."

We have added the word "physical" as suggested.

P5, l3: Also mention the amplifier which becomes important later on.

The whole radiometer plate, which includes the mixer and amplifiers is temperature-controlled. Hence, the amplifiers are also listed now.

P6, l2: "..doubled twice, POTENTIALLY allowing …Here only the standard set… is used." In the beginning I was confused here so this might also help others.

We have made the suggested changes in the manuscript.

P6l25: I find "half-hemisphere" =quarter sphere rather confusing. It is not necessary to mention it but rather it is important to say what "Both functions" in the next sentence means.

We have changed the sentences to: "… to the different directions it is pointing towards. Both, the filter functions, as well as the antenna diagram, have been measured …".

P6/5: I don't think that the logic for the second part of the sentence is correct "despite the fundamental assumption in MTP calibration that this relation is always linear ". I suggest to change to "Amplifiers might change their characteristics and thus the relation between the recorded signal and the source temperature, i.e. the calibration parameters, change." You still assume linearity. For nonlinearity either higher order terms or an exponential relationship would be need to be considered in the calibration equation and determined by introducing a third reference point.

We agree that the suggested sentence is clearer, and have made the according change in the manuscript.

P7,l4: noise diode signal strength

"strength" was added after "noise diode signal".

P7,l24: "The gap in the centre is due to the use of a double-side-band receiver as explained in Sec?"

The suggested change has been made, including a reference to Section 2.

P8,l3: Anntenna pattern: "It is actually mainly defined by the shape of the rotating mirror at the front of the instrument." The text did not provide any information on the shape and I can not see something in Fig. 1. If it is a flat mirror then only the horn is important. If it is used for beam forming, e.g. an offset parabolic mirror, it is important to mention that? Has the antenna beamwidth not been calculated using gaussian beam optics?

In Table 1 we explicitly mentioned that the rotatable mirror is used to define the beam shape. We have now added that it is hyperbolic. The beam shape was measured in the laboratory, as described in the manuscript, and the parameters stated in the text are derived from that measurement, confirming the numbers provided by the manufacturer.

P9,l11: "itself IS monitored"

We have made the suggested change.

P9,l11: Over this large temperature range it is difficult to see by eye if a linear relation ship fits the data better than a non linear/exponential one? This comment is maybe a bit picky but one could easily check the significance of the statement. What I further notice is that the slope for the dark blue load seams to be lower than the one of the light blue load for the lower two channels and vice versa for the upper channel. As this is shown in counts rather than in brightness temperatures it is difficult to judge if this is significant. Why don't the authors use a simple/standard calibration that would help the reader to judge the impact – this could just be plotted on the second y-axis for illustration – of course an offset needs to be introduced to separate the channels? This comment also holds for the following two figures as nobody is interested in counts,

This comment relates to figure 5, which illustrates the direct linear relationship between the measured signal (given in counts) and the source temperature (ambient targets at different positions within the cold chamber). A conversion to brightness temperatures already uses a calibration calculation based on some standard values, which – as is shown later in the manuscript – clearly depend on the current cold-chamber temperature, and corresponding instrument state. Hence a conversion to brightness temperatures may lead to a wrong interpretation of data. This, along with the other findings that calibration coefficients differ for different frequencies, has led to the decision to plot the raw signal in counts, and we would refrain from changing the figures in the manuscript.
When performing quadratic or even cubed fits to the data, the parameters clearly indicate that only the linear term contributes significantly (fit coefficients at least three orders of magnitude larger than the quadratic or cubed term). We have added a sentence in the manuscript reflecting this.
On the comment concerning the differences between the dark and the light blue lines: Please note that due to the limited space available, one of the targets is placed closer to the ventilation of the chamber than the other, so the two ambient targets do not have the exact same temperature!

P9,l30 – same comment relate counts to Tb

The sentence was changed to: " … can be characterised by a Gaussian distribution with a 30 standard deviation of approximately 6 cnts (approx. 0.25 K) and the mean at 0 cnts.

P10, l11: In addition to the term "noise figure" the use of "long-term stability" is unclear. Suggest: "This is strong evidence that the HALO-MTP noise characteristics do not change between flights, and the laboratory assessment can be used to characterize measurements of different campaigns and serve as information for the retrieval development."

The sentence was changed to: "This is strong evidence that the HALO-MTP noise characteristics do not change between flights, and the laboratory assessment can be used to determine overall instrument-health and comparability of measurements in between campaigns, also serving as information for the retrieval development."

P12, l20: In fact if you are flying high (thus having an optical thin atmosphere above you) and assume horizontal homogeneity you could do the tipping curve calibration from the different elevation angle measurements pointing upwards (probing different opacities). For sure, I don't expect that you do that but it is not the instrument design limiting you.

While according to Han and Westwater, 2000 the antenna beam width is a major source of uncertainty in this calibration approach, we agree that there are other limiting factors, too.

The sentence has been changed to: Other calibration methods, such as the tipping curve calibration (e.g. Küchler et al., 2016, or Han and Westwater, 2000) are not available for the DLR MTP due to the given instrument design (mainly antenna beam width), potentially fast-changing atmospheric conditions, due to the moving platform, influencing radiative transfer calculations needed in this approach, as well as the need for an efficient measurement strategy.

The reference to Han and Westwater, 2000 was added in the reference list.

P14 l7: Please again relate counts to TB

We added that 20 counts correspond to approximately 0.83 K (using the reference calibration parameters from Table 7).

P15, l18 – suggest to say absolute accuracy

We have made the suggested change.

P15, l29: "The accuracy of the temperatures " – be more specific. Which T, which type of accuracy?

We have added the word "brightness" to make clear, that brightness temperatures are meant.

P16, l17: Eq. 4.5 give the noise equivalent temperature NeDT not the measurement uncertainty. This is not the variance (rather square root of variance) of the measurement noise but the radiometric resolution (see Eq. 6.13 by Woodhouse, 2017) and can be estimated as standard deviation from time series, e.g. from Fig.6 I would estimate something like 0.3 K. This already roughly corresponds to the values in Fig. 11. Note, the effect of the

passband is negligible compared to other deviations from an ideal radiometer and should not be mentioned.

Equation 6.52 in Ulaby et al., 1981, which is the one stated in the Manuscript, is derived from the definition of the variance. We have changed the equation in the manuscript to look the same as in Ulaby et al., 1981, and used the term ΔT_{theo} in the text for continuity. As explained in the end of the paragraph, this theoretical equation does not consider any gain variations.
The gain variation is still part of the calculation of the standard deviation (Figure 7), since we only subtracted the linear fit (drift of the instrument), not the running average (blue line in Figure 7). Only that would have included the gain variations, and would have led to a similar value as the theoretical one. However, that value would not be representative of the real flight situations, where changes of the atmospheric temperature are frequently observed, and must remain detectable in the MTP data.

Maybe it is good to clarify that in its simplest concept measurement uncertainty includes systematic errors (bias, absolute accuracy) and random errors (noise, precision).

We have added a sentence in the beginning of the section stating: “The overall uncertainty of the MTP measurements includes both systematic errors (e.g. the bias, described above), and random errors (e.g. noise). In case of the HALO MTP, the former can be related to HALO TS, as shown above, the latter mainly influences measurement precision, and influences the ability of the instrument to pick up atmospheric temperature fluctuations. In the literature (e.g. Ulaby et al., 1981), …”

P18, l14: can you clarify whether this is random or systematic?

The deviation is mainly a systematic error (the biggest source being the aircraft pitch, which is now stated in the manuscript); however, it is dependent on the flight conditions, as the offset changes with speed and altitude of the aircraft. Possible vibrations may add a random error contribution. Since there is not enough data to make statistically significant evaluations of the offset, we cannot say more than is already stated in the manuscript.

P18, l26: It is likely only relevant if hydrometeors or WV variations occur,,

This comment refers to the discussion of synthesizer errors. We agree that those are not a large error source, as discussed in the manuscript. It is clear that the frequency-dependent influence of water vapour would only be noticeable, if strong variations in the frequency would occur. The argument stated in the text is still valid: the synthesizer error must be large enough to cut out a portion of the Oxygen absorption line to make the changed influence of water vapour absorption noticeable. In that case, however, probably the whole radiative transfer calculation must be adapted anyway, and the health of the instrument should be checked.

P20, l7: please transfer to Tb and discus in respect to the later results and the term precision. It is especially important here as you want to give an idea which amplitude of gravity waves (be clear on that) can be detected.

6 counts correspond to slightly less than 0.25 K, which was added in the text. We also added a half-sentence at the end of the paragraph, so that the last sentence now reads:
“ […] the presented characterisation of the HALO-MTP noise allows the identification of significant atmospheric signals in MTP measurement time series, as long as the amplitude of

the atmospheric signal is larger than the measurement precision (i.e. significantly larger than 0.38 K), and wave-like signals (e.g. caused by gravity waves) can be clearly separated from the noise-induced structures, caused by the auto-correlation of MTP measurements.”.

P20l20: “To achieve this accuracy, the necessity of an offset-correction relative to HALO TS” Here you talked about precision not about the bias. Just delete the first part “To achieve this accuracy”

We have made the suggested change in the manuscript.

P20, l30. “clearly dominated by the contribution from measurement noise.” Is this really true if I look at Fig. 10.

In the manuscript, the following sentence explains that other potential sources of uncertainty for brightness temperatures (i.e. the retrieval input) do not have as much influence on the overall uncertainty than other potential sources of uncertainty. Potential drifts seen in Figure 10 (which depicts data recorded during a measurement flight, not in the laboratory), are most likely caused by changing atmospheric conditions. We remind the reviewer that the offset-correction only uses one single value; the leg-mean HALO-TS temperature. It is important that atmospheric temperature changes occurring during a flight leg remain detectable, as seen in Figure 10. We have added this in the manuscript text (both, in Section 4, as well as in the Summary).

P21, l12: “all necessary instrument parameters” I wouldn’ be so confident!

The sentence was changed to: “Overall, this study summarises the investigation of instrument parameters and characteristics, necessary to accurately analyse…”
It was clearly stated in our study that there are limitations to what can be characterised in this case, and we have performed those characterisations that are necessary and helpful to understand the recorded data (i.e. being aware of uncertainties in data time series, and knowing the abilities of the instrument). Thus, we are confident, that all necessary information is provided in this study to allow for correct data interpretation and evaluation. At the same time, we do acknowledge the reviewer’s comment from the beginning; that – as this is the nature of scientific progress – the future might show that there might be some detail that needs to be considered, of which we cannot be aware of at this point.

Measurement Characteristics of an Airborne Microwave Temperature Profiler (MTP)

Mareike ~~Kenntner[1]~~Heckl[1], Andreas Fix[1], Matthias Jirousek[2], Franz Schreier[3], Jian Xu[3], Markus Rapp[1]

[1] Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany
[2] Deutsches Zentrum für Luft- und Raumfahrt, Institut für Hochfrequenztechnik, Oberpfaffenhofen, Germany
[3] Deutsches Zentrum für Luft- und Raumfahrt, Institut für Methodik der Fernerkundung, Oberpfaffenhofen, Germany

Correspondence to: Mareike ~~Kenntner~~Heckl (Mareike.~~Kenntner~~Heckl@DLR.de)

Abstract. The Microwave Temperature Profiler (MTP), an airborne passive microwave radiometer, measures radiances, recorded as counts and calibrated to brightness temperatures, in order to estimate temperature profiles around flight altitude. From these data, quantities such as potential temperature gradients and static stability, indicating the state of the atmosphere, can be derived and used to assess important dynamical processes (e.g. gravity waves or stability assessments). DLR has acquired a copy of the MTP from NASA-JPL, which was designed as a wing-canister instrument and is deployed on the German research aircraft HALO. For this instrument a thorough analysis of instrument characteristics has been made in order to correctly determine the accuracy and precision of MTP measurements.

Using a laboratory set-up, the frequency response function and antenna diagram of the instrument ~~was~~were carefully characterised. A cold-chamber was used to simulate the changing in-flight conditions and to derive noise characteristics as well as reliable calibration parameters for brightness temperature calculations, which are compared to those calculated from campaign data.

The MTP shows quite large changes of the instrument state, imposing considerable changes in calibration parameters over the course of a single measurement flight; using a built-in heated target for calibration may yield large errors in brightness temperatures, due to a misinterpretation of the measured absolute temperature. ~~By applying~~Applying herein presented corrections to the calibration parameter calculations, the measurement noise becomes the dominant source of uncertainty and it is possible to measure the brightness temperatures around flight level (closely related to the absolute temperature close to the instrument) with a precision of 0.38 K. Furthermore, radiative transfer ~~calculations~~simulations, using the ~~radiative transfer model~~ Py4CAtS package in a pencil-beam approach, indicate that the ~~vertical observation~~altitude range of sensitivity of the MTP instrument can be increased by applying a modified measurement strategy.

This is the first time such an extensive characterisation of ~~a~~an MTP instrument, including a thorough calibration strategy assessment, is published. The presented results, relevant for the wing-canister design of the MTP instrument, are important when processing MTP data: Knowledge of the

relevant uncertainties and instrument characteristics is essential for retrieval setup and mandatory to correctly identify and interpret significant atmospheric temperature ~~fluctuation~~fluctuations.

1 Introduction

Aircraft campaigns have long been used to study atmospheric composition and dynamics. ~~One~~Here, one important variable to be determined ~~during aircraft measurements~~ is the atmospheric temperature, ideally not only at flight level, as provided in high resolution by the standard aircraft instrumentation. For this measurement it is desirable to use a remote sensing technique, which provides good horizontal and vertical resolution. A variety of instruments and techniques exists, many of them used in ground-based set-ups or installed on satellites. For aircraft instruments, the line-of-sight is always an important factor, as well as the ability to record data fast (providing high horizontal resolution). It is desirable to use a robust instrument design, able to perform despite frequent changes of conditions due to flight patterns and geographical region of deployment. On the German High Altitude LOng range research aircraft (HALO; Krautstrunk and Giez, 2012), the microwave temperature profiler (MTP; Denning et al. 1989) complements other instruments such as the ~~Basis~~Basic HALO Measurement and Sensor System (BAHAMAS) that measures the temperature amongst other parameters at flight level, or dropsondes. In contrast to such in-situ instruments, the MTP scans through the atmosphere at different viewing directions, providing temperature profile information at, above, and below flight level. A copy of this compact wing-canister instrument, which was originally designed by NASA-JPL, has been transferred to DLR, and was modified and certified for operation on HALO. On that aircraft it constitutes a valuable addition to the scientific payload as the data recorded by the MTP facilitate the interpretation of trace gas measurements taken during flight (e.g. by indicating tropopause height and static stability) and increase the atmospheric region on which information can be gathered. Combining MTP and dropsonde data (e.g. for cross-validation), or exploiting the synergy with the airborne multi-wavelength water ~~vapor~~vapour differential absorption lidar (WALES; Wirth et al., 2009) ~~offer~~offers the opportunity to increase the insight into atmospheric processes targeted during measurement flights. Its observation range at, above, and below flight level plus its small size and weight clearly sets the MTP apart from the HALO Microwave Package (HAMP; Mech et al., 2014), which is another optional instrument package deployable in HALO's bellypod capable of retrieving both humidity and temperature profiles, as well as the liquid water path (Jacobs et al., 2019), below flight level by means of passive microwave radiometry.

The value of MTP data is also visible through its continued use in many aircraft campaigns. Since its invention in the late ~~1970ies~~1970s, the MTP has been deployed in a number of aircraft campaigns (Mahoney and Denning, 2009), and continues to be developed to meet today's standards of technical requirements~~,~~ and data recording. In the past, MTP data has been used to interpret in situ measurements of trace gases (e.g. Marcy et al.,

2007; Thornton et al., 2007; Spinei et al., 2015), aerosols, (e.g. Gamblin et al., 2006; Popp et al., 2006; Schwarz et al., 2008), and to assist the study of cloud physics (e.g. Corti et al., 2008; Jensen et al., 2010; Schumann et al., 2017; Urbanek et al., 2017), and dynamics in the atmosphere (e.g. Tuck et al., 1997, 2003; Dörnbrack et al., 2002; Sitnikova et al., 2009). Other studies, focussing exclusively on MTP data, include the derivation of the boundary layer height from MTP potential temperature isentropes (Nielsen-Gammon et al., 2008), investigation of mixing processes within the polar vortex (Hartmann et al., 1989), or of the cold point temperature and mesoscale temperature fluctuations, derived as the difference to the mission average temperature, in the upper troposphere and lower stratosphere (UTLS), in connection to tropical weather disturbances~~.~~ (Davis et al., 2014). Furthermore, MTP measurements have been utilised to investigate gravity waves in the atmosphere. Studies focussed on general overviews (Gary, 2006, 2008), the formation of polar stratospheric clouds (PSCs; Murphy and Gary, 1995 and Tabazadeh et al., 1996), or the characterisation of gravity waves encountered during flight (Gary, 1989, Chan et al., 1993, Dean-Day et al., 1998, and Wang et al., 2006~~-~~). Based on these mesoscale temperature fluctuation analyses, a number of modelling studies~~ have been published,~~ aimed at improving the understanding and numerical description of atmospheric gravity waves, including studies by Bacmeister et al. (1990, 1996, 1999), Pfister et al. (1993), Cho et al. (1999), Leutbecher and Volkert (2000), Dörnbrack et al. (2002), and Eckermann et al. (2006). Especially for studies focusing on mesoscale temperature fluctuations or vertical temperature gradients, precise knowledge of the instrument characteristics, such as intrinsic noise~~,~~ and the precision of the measurements is necessary, e.g. when identifying potential gravity wave signals within the time series of MTP data. Knowing the true range of sensitivity is also necessary to understand the shape and characteristic structures within the retrieved temperature profiles.

Despite the continuous use of data from various MTP instruments in many studies over the past decades, a thorough instrument characterisation and ~~derivation~~estimation of measurement accuracy (i.e. the deviation from the true value; mostly influenced by systematic errors), and precision (i.e. the spread of the individual measurements; mostly influenced by random errors, such as measurement noise), has not yet been published. For the first time, this study presents ~~all~~a thorough investigation of relevant instrument characteristics of the HALO-MTP instrument. All measurements shown in the following sections are important to correctly choose retrieval settings and interpret time-series of MTP data. They were conducted without disassembling the instrument, granting that the hardware characteristics are comparable to previous mission deployments, and guaranteeing the continued airworthiness of the instrument on the HALO aircraft. Knowing the instrument characteristics is the foundation for correct analysis and interpretation of data recorded by the HALO-MTP. The following sections present a brief description of the instrument and its measuring principle (Section 2), measurements of the instrument response function, the antenna diagram, and other inherent characteristics, such as measurement noise (Section 3), a discussion of calibration strategies to determine the best practice (Section 4), including a discussion of the influence of flight level changes on the instrument state and measurement performance. Further error sources, and some discussion of possible improvements to the measurement strategy are presented in Section 5. The findings are summarised in Section 6.

2 Instrument description

The first MTP instrument was developed in the late 1970s by Bruce Gary and Richard Denning at the Jet Propulsion Laboratory (NASA-JPL) for research on clear air turbulence (CAT; Gary, 1989). Since its first deployment in the Stratospheric-Tropospheric Exchange Project (STEP) in Australia, 1987, the MTP has widely been regarded as an instrument providing valuable background information on the state of the atmosphere, and several instrument designs have been realised. The latest development is the MTP as wing-canister instrument (see Figure 1, left panel), which can be mounted underneath the wing of a research aircraft (e.g. Haggerty et al., 2014). Two such instruments were built. One has been deployed on the NCAR GV since 2008 (e.g. Lim et al., 2013; Davis et al., 2014; Haggerty et al., 2014), the other (hereafter referred to as HALO-MTP) was acquired by DLR, and has been flown on the German research aircraft HALO (Krautstrunk and Giez, 2012). This design was first introduced in 2008. Details of the instrument design and some discussion on standard measurement settings can be found in Mahoney and Denning (2009) and Lim et al. (2013).

In the following, an overview of the characteristics of this specific version of MTP instruments is given, starting with a brief introduction of the measurement principle and followed by a description of specific radiometric hardware (see also Table 1), built into the HALO-MTP. All results of the following sections are representative for this specific MTP instrument design.

Measurement principle

The concept of measurements of the MTP as a passive total-power radiometer (Denning et al., 1989; Ulaby et al., 1981) is straightforward. The MTP records thermal radiation mainly emitted by oxygen molecules in the atmosphere. As many radiometers measuring atmospheric temperature, the MTP uses absorption lines of the 60 GHz oxygen complex ('V-band'), which are caused by magnetic dipole transitions (Liebe et al., 1992). Passive radiometers pick up the energy transported by the photons emitted in these transitions. In this part of the spectrum, the Rayleigh-Jeans relation (e.g. Ulaby et al., 1981) can be used to describe the source function of the radiance picked up by the MTP:

$$B(\nu,T) = \frac{2h\nu^3}{c^2} \cdot \frac{1}{\exp\left(\frac{h\nu}{k_B T}\right) - 1} \cong 2\frac{\nu^2}{c^2} \cdot k_B T \quad (Eq.\,2.1)$$

~~implying~~in which the measured radiance is equal to the brightness B of a blackbody described by Planck's law. This equation implies a linear relationship between the measured radiance ~~B~~ and the temperature T of the black body source at a certain frequency ν, using the Planck-constant, h, velocity of light, c, and the Boltzmann-constant k_B. This temperature, T, is referred to as brightness temperature (TB), which is the temperature of on an ideal blackbody, emitting the equivalent radiance. The recorded TBs have to be converted to absolute temperature profiles by using a

retrieval algorithm that utilises forward radiative transfer calculations~~.~~, in which, ideally, all possible impacts on the measured radiance (e.g. water vapour, nitrous oxide, or hydrometeors, etc.) have to be considered. For correct interpretation of the fluctuations found in the retrieved temperature fields, it is necessary to have precise knowledge of the instrument characteristics, such as the instrument response function, antenna diagram (see Section 3), and the precision and accuracy of the brightness temperature measurements that are input to the retrieval algorithm (see Section 4).

In the MTP instrument, a horn antenna ~~is used as the receiver for~~guides the incoming atmospheric radiation and, together with the hyperbolic shape of the rotating mirror, determines the spatial response function. The measurements are based on the heterodyne principle, which means that through mixing with a defined frequency, the local oscillator frequency (LO), the incoming signal is converted to ~~a~~ an interim frequency (IF) near base-band. ~~Then, by low pass filtering only the part of the incoming radiation spectrum around the current LO is selected. The MTP uses a double-sideband receiver. Thus,~~Both difference frequencies below and above the LO frequency are down-converted to the ~~same base~~ IF in the double side band~~, so~~ receiver. Low pass filtering supresses any incoming radiation outside the IF bandwidth of 200 MHz such that ~~both flanks~~the symmetric spectrum around the current LO frequency is measured with only a minor gap of ~~an oxygen absorption line can be further processed. The passing~~approx. 20 MHz at the LO frequency. The IF signal is converted to a voltage, which is proportional to the squared input amplitude, representing the power. This voltage is finally translated to a digital count number, stored in the MTP data file, and later converted into a brightness temperature through calibration (see Section 4). ~~-~~The physical temperatures of the important radiometric parts of the radiometer, such as the mixer, synthesizer, amplifiers, as well as the electronics are stabilised, to minimise the influence of the changing conditions during a research flight on the instrument state, and to protect the electronic parts from malfunction due to condensation (see Mahoney and Denning, 2009, for further details).

Using a rotating mirror in front of the instrument's feed antenna (number 3 in right panel of Figure 1), the direction from which the radiation is collected can be changed. Moving through a single set of elevation angles, as well as the set of frequency channels at each of those elevation angles, is referred to in the following as a measurement "cycle". This procedure enhances the vertical resolution of measurements in comparison to non-scanning measurements which derive the altitude information solely from exploiting frequency-dependent differences in optical depth of the atmosphere. The MTP combines both principles: In its standard deployment settings, as programmed in the original JPL instrument software, ten viewing angles are being used during one measurement cycle; five above the horizon, four below, and one pointing exactly towards the horizon. At each angle, measurements at three frequency channels, corresponding to the frequencies of three strong oxygen absorption lines, are subsequently performed (see Table 1), before moving to the next elevation.

Furthermore, a calibration target is built into the instrument (number 2 in right panel of Figure 1), to which the mirror points after each cycle of atmospheric measurements. The target itself consists of carbon-ferrite on an aluminium plate, which is heated to a constant temperature (approximately 45 °C) using two conventional power resistors. The calibration target is surrounded by a 1-inch-thick Styrofoam insulation which is transparent for microwave radiation. The signals recorded while pointing towards the heated target are combined with a noise diode (ND) signal and used for calibration (see Section 4) to convert the measured signal to a TB, which is usually not equal to the outside air temperature, since the measured signal is influenced by multiple layers of the atmosphere. To derive absolute temperature from the radiation measurement, radiative transfer calculations have to be carried out and compared to the measured radiances by applying a retrieval algorithm in post-processing. The instrument characteristics presented in the following sections of this work all correspond to the raw measurements or the brightness temperatures, which are input to such a retrieval algorithm. However, the discussion of retrieval algorithms and related uncertainties are beyond the scope of this paper.

Specific wing-canister instrument hardware characteristics

Differences to older instrument designs are presented in Lim et al (2013) and Haggerty et al. (2014). The most important upgrade is that the LO is now defined as a frequency near (or ideally at) an oxygen absorption line centre so that the two flanks that are measured belong to the same line, which lowers the brightness temperature error, as discussed in Mahoney and Denning, 2009, and Lim et al., 2013. The instrument is pointing forward, measuring the temperatures of air masses in front of the aircraft. The filter bandwidth of the HALO-MTP is fixed to ±200 MHz around the LO, with a gap of approx. 20 MHz at the line centre (See Figure 2, left panel). The synthesizer used to generate the LO for down-conversion of the signal can be tuned between 12 GHz and 16 GHz. The synthesizer output is doubled twice, potentially allowing for a frequency range of 48 GHz to 64 GHz for atmospheric measurements. The pre-set ("standard") set of used LO frequencies was chosen based on the considerations presented in Mahoney and Denning, 2009, and Lim et al., 2013. This set of frequencies is used for this study.

Two significant modifications to the original instrument were made by DLR: An embedded computer and an inertial measurement system including a Global Positioning System (GPS) antenna. In the original set-up a Visual Basic software package was provided by NASA-JPL to run the instrument during research flights. With the on-board computer and integration of the inertial sensor, this software was translated to a LabView code, which was adjusted to use the additional data provided by the inertial sensor. With those modifications, the HALO-MTP can be operated autonomously, i.e. independent from a connection to a cabin computer, which is still provided, and can be used, e.g. to adjust settings during research flights using the HALO LAN network.

The HALO-MTP was first deployed during the Midlatitude Cirrus Experiment (ML CIRRUS) in 2014 (Voigt et al., 2017). The focus of this mission was to probe natural cirrus clouds as well as contrail cirrus throughout various stages of their life-cycles. The MTP was part of the wing-probe instrumentation and recorded data during all mission flights. In total, the HALO-MTP produced almost 63 hours of data during 13 mission flights, recording 17476 individual measurement cycles. Data from this campaign will be used to derive the HALO-MTP noise ~~figure~~characteristics and in the investigation of calibration methods in the following sections.

3 Characteristics of the wing-canister MTP flown on HALO

To retrieve absolute temperature profiles from the MTP measurements, radiative transfer calculations are carried out, to model the measured radiance in a defined atmospheric state. To correctly do so, the instrument transmission function (c.f. Figure 2, left panel) has to be known. This function is defined by the instrument's filter function, which determines which part of the recorded spectrum is used in data processing. Moreover, the antenna diagram (c.f. Figure 2) shows how sensitive the receiver is to the different directions ~~in the half-hemisphere~~ it is pointing towards. Both, the filter functions, as well as the antenna diagram, have been measured in a stable laboratory environment (Section 3.1).

Since this MTP instrument is mounted to the outside of the aircraft (c.f. Figure 1, left), the instrument experiences changes in surrounding pressure and temperature during measurement flights in which flight level changes can be quite common. During a single mission flight, the air temperature surrounding the aircraft can change from around 300 K on the ground to as low as 190 K in the tropopause. These changing ambient temperatures of the MTP can influence the performance of the instrument: Amplifiers ~~may~~might change their characteristics and thus the relation between the recorded signal and the source temperature, ~~despite~~i.e. the ~~fundamental assumption in MTP~~ calibration ~~that this relation is always linear.~~parameters, change. Moreover, the noise diode used for calibration may change its signal strength (see Section 4.2), and the overall instrument noise can be affected. However, the noise characterisation is particularly important when interpreting temperature fluctuations in a time series of MTP data. Knowing possible periodicity in the noise signal is essential to distinguish between real periodic atmospheric temperature fluctuations, e.g. those caused by gravity waves, and instrument noise. To test all of these characteristics, the MTP was placed inside a temperature chamber (Figure 3), to simulate the changing outside air temperature during mission flights. In these tests, the influence of the changing surrounding temperature on the linearity of the sensor is tested (Section 3.2), and the measurements are used to determine the noise characteristics of the HALO-MTP. The laboratory results are also compared to the noise characteristics derived from airborne ML CIRRUS mission data (Section 3.3). Possible calibration strategies and the effect of changing instrument state on the calibration of data will be discussed in Section 4.

3.1 Instrument function

The measurements of the instrument transmission functions, as well as of the antenna diagram, were made in a chamber completely covered in microwave absorbers. The MTP was installed on a rotatable platform. A tuneable signal source with a horn antenna was placed in 5-m distance to the MTP instrument. The signal was then measured by the MTP, as well as by a power meter for reference. The power of the source signal was chosen such that the MTP signal was well above its inherent noise level. For the measurement of the filter function, the source frequency was tuned between LO – 300 MHz to LO + 300 MHz in steps of 1 MHz. The measured signal is normalised and then corrected for frequency dependency, based ~~on~~onthe Friis Transmission Equation (e.g. Balanis, 1997):

$$cnts_{corr} = \frac{cnts_{\mathrm{norm}}}{f^2/(min(f))^2} \quad (Eq. 3.1)$$

Finally, the signal power of the source, $P_{\mathrm{corr}}(f)$, is taken into account in a final normalised signal representing the relative forward transmission, $cnts_{\mathrm{final}}$:

$$cnts_{\mathrm{final}} = \frac{cnts_{\mathrm{corr}}}{P_{\mathrm{corr}}(f)} \quad (Eq\ 3.2)$$

The resulting instrument transmission functions for the three standard frequency channels are shown in Figure 2, left panel. It shows symmetrical shapes for all frequency channels' functions (i.e. radiances are recorded symmetrically from both flanks of the probed oxygen line), confirming a transmission of the signal between ±200 MHz around the LO (width of the plateau). The gap in the centre is created by the receiver architecture, using a double-side-band ~~biased mixer.~~receiver as explained in Section 2. A certain 'waviness' with an amplitude of about 0.5 dB is visible next to this gap. To exclude reflections from the chamber as a source, the measurements were repeated multiple times with slightly different positioning of the source antenna and the instrument. Since the results were similar in all measurements, the source of this 'waviness' is attributed to some internal source within the instrument, due to electromagnetic wave propagation through the instrument parts.

The main result of measuring the antenna diagram is the field-of-view (FOV) of the instrument, defined by the full width half maximum (FWHM; red, dashed lines in Fig. 2 middle and right panel) of the antenna diagrams. It is actually mainly defined by the hyperbolic shape of the rotating mirror at the front of the instrument. The measurement was made using the same laboratory setup as for the measurement of the transmission function. Both the horizontal and the vertical plane were measured in steps of 1° rotation. The symmetric shape of the diagram implies that radiance is picked up equally strong from all directions. Note that the maxima of the side-lobes in the antenna diagrams have a maximum at -30 dB, meaning the signals from these spatial directions are 1000 times weaker than the signal picked up from the main viewing direction. The FOV is about 7.0°-7.5° in the horizontal and about 6.5°-7.0° in the vertical at all frequencies.

A spill-over measurement of the horn antenna (as investigated, for example in McGrath and Hewison, 2001) as well as a test of the stability of the LO frequency generator was not possible, as this would have required disassembling parts of the instrument, which was not an option at that time due to aircraft certification issues.

3.2 Temperature dependence of MTP characteristics

To investigate the temperature-dependence of instrument performance, a series of measurements inside a cold chamber was performed (see Figure 3). During this measurement series, the temperature of the cold chamber was successively lowered from 21 °C to -15 °C in steps of 5 °C. This temperature range resembles the temperatures the MTP experienced during its deployment in the ML CIRRUS campaign in 2014, as shown in Figure 4, right panel. The pod air temperature sensor monitors the temperature inside the MTP's housing during the flight (c.f. Figure 1, right). In the cold chamber, the housing was not installed to prevent over-heating of the instrument at higher temperatures. As a result, the readings of this sensor show the air temperature inside the cold chamber. The scanning unit temperature sensor keeps track of the temperature of the MTP instrument within close proximity to the crucial parts of the radiometer, such as the amplifiers or the mixer. The readings of this sensor give an impression of the state of the instrument and its thermal stability. It can be seen that the response to lowering the cold chamber temperature is different between the two sensors. This is caused by the placement of the sensors, one being closer to some heated parts of the instrument, indicating that changes in the environment of the instrument are not equally influencing all parts of the instrument. Moreover, from the readings of the scanning unit temperature sensor (black line in left panel of Fig. 4) it can be seen that the MTP instrument takes some time to stabilise under the new temperature conditions. This time required for stabilisation depends a lot on the operating environment, such as size of the laboratory space, ventilation, or ambient temperature. In this setting, it takes up to 15 minutes after the initial temperature change. Only those parts of the measurement series are used in which the scanning unit temperature is stable (the difference between two readings being smaller than an empirical threshold value of 0.04 K), to exclude effects from the instrument adjusting to new environmental conditions.

Along with the MTP instrument two microwave absorbers (Telemeter Electronic GmBH EPP51 broad-band pyramidal absorber) at ambient temperature (hereafter referred to as 'ambient targets'), and a similar microwave absorber submerged in liquid nitrogen (hereafter referred to as 'cold target') were placed in the chamber, in order to perform calibration measurements throughout the complete measurement series. The third type of calibration target used in this measurement series is the built-in calibration target of the MTP instrument (see Section 2), hereafter referred to as 'hot target'.

Linearity of the sensor

Using the measurements of the two ambient targets installed within the chamber, it can be shown that for the HALO-MTP the linear relation between the source temperature and the measurement output is given at all standard frequency channels (see Fig. 5). Fits to the data using higher orders, show that only the liner term significantly contributes to the data fit. Since not only the temperature of the target changed during this test, but also the temperature of the sensor unit itself is monitored (see Fig. 4), it can also be established that the linear relationship between the measured signal and the source temperature is maintained throughout changing conditions. The measurements corresponding to the two individual ambient targets (different line colours in Fig. 5) are nearly identical, proving consistency of measurements.

The calibration parameters needed to calculate the brightness temperature (T_B) from the measured signal ('counts'; c) are therefore the y-intercept (receiver noise temperature; T_R), and the slope of the line (s_{cal}), drawn through two points defined through measurements of calibration targets at known temperatures:

$$T_B = c \cdot s_{cal} - T_R \quad (Eq. 3.3)$$

In Section 4 it will be shown that those parameters are depending on the instrument state, and can be related to housekeeping data. Please note, that in the classical microwave notation, the calibration is actually defined inversely as $T_B = c \cdot G + T_R$, in which the gain (G) is equal to the inverse slope as defined in the above equations.

3.3 Noise characterisation

The instrument's noise ~~figure~~ was characterised using the signal measured when pointing towards the hot target. It is assumed that due to the temperature stabilisation of the target, the mean measurement signal should not change over time. Hence, the deviation from the mean represents the noise added by the instrument. An example of the measured signal while looking at the hot target during one measurement segment at constant cold-chamber temperature is shown as the grey line in Fig. 6. Obviously, absolute stability can hardly be reached in a cold environment, while parts of the sensor unit are heated to approximately 40°C, which is taken into account by applying a linear fit to the measured data of one segment (black line in Fig. 6) instead of simply subtracting the mean (blue line in Fig. 6). Using all segments of the cold-chamber measurements, the resulting HALO-MTP noise ~~figure~~characteristics, as shown in Figure 7 (top), can be ~~characterised~~described by a Gaussian distribution with a standard deviation of approximately 6 cnts (approx. 0.25 K) and the mean at 0 cnts.

The same method as for the cold chamber measurements is used for HALO-MTP data recorded during the ML CIRRUS campaign in 2014. Here, the criterion used to determine flight segments with nearly stable instrument states is a difference of the scanning unit temperature of less than 0.04

K between two measurement cycles. Additionally, it was made sure that no altitude changes were made ($\Delta z \leq 25$ m) or curves were flown during these segments. From all ML CIRRUS mission and test flights, 61 segments could be identified that satisfied the criteria and were at least 5 minutes long, to ensure significance of statistical results from the length of one segment: With the length of one measurement cycle (including a calibration measurement at the end) at 13 s each segment includes at least 22 recordings of the hot target measurement signal. The middle panel of Figure 7 shows the noise characteristics at 56.363 GHz. The results from the flight data evaluation are in excellent agreement with the values found in the laboratory environment, showing even smaller standard deviations of 5.2-5.7 cnts, depending on the frequency channel. This is strong evidence that the HALO-MTP noise ~~figure does~~characteristics do not change between flights, and the laboratory ~~characterisation~~assessment can be used to determine ~~long-term stability of the~~ overall instrument-health and comparability of measurements in between campaigns, also serving as information for the retrieval development.

For the spectral analysis of the MTP measurement noise ~~figure~~ the 61 ML CIRRUS mission flight segments are used again. Due to the varying lengths of the individual flight legs, the data is concatenated to a single time line for spectral analysis. The power spectrum of the noise signal of the HALO-MTP at 56.363 GHz, as shown in Fig. 7 (bottom), reveals that the measurement noise can best be described as a red noise, which is characterised by the auto-correlation α between a data point of the time series and its precursors. According to Torrence and Compo (1998), the corresponding theoretical noise power spectrum for a range of wave numbers k, P_k, is given by:

$$\text{~~P~~}_k P_k = \frac{1-\alpha^2}{1+\alpha^2-2\alpha\cos(2\pi k\,/\,N)} \qquad (Eq.\,3.4)$$

For the three standard frequency channels, the lag-1 autocorrelation of MTP measurements during the ML CIRRUS campaign is $\alpha \cong 0.7$. Fit parameters characterising the noise ~~figure~~characteristics at the three standard frequency channels are summarised in Table 2.

With the above findings, characterising the HALO-MTP measurement noise ~~figure~~ as Gaussian-shaped, with mean at 0 counts, and a standard deviation of 6 cnts, as well as with the knowledge of the inherent periodic structure of the noise signal, it is now possible to determine whether periodic structures in a MTP temperature measurement time series are significant (high probability that they result from atmospheric temperature fluctuations), or noise-induced. Additionally, the standard deviation of the Gaussian distribution of noise values can be used to determine the variance of TBs derived from the raw signals, once the calibration parameters are known. While it would be interesting to investigate the integration time needed to significantly reduce the measurement noise (e.g. by applying the Allen variance), any increase in integration time will decrease the horizontal resolution of measurements taken on a jet-engine aircraft. The current settings present a well-chosen compromise between the measurement noise and the horizontal resolution of the MTP data (Mahoney and Denning, private comm.). As the calibration measurement is

already made at the end of each measurement cycle, the instrument is calibrated as frequently as possible, and the calibration performance cannot be further improved by adding more calibration measurements during mission flights.

4 Investigation of calibration methods for the HALO-MTP

In Section 3.2 it was shown that there is a linear response in the measured signal to changes in the source temperature, so that the measured signal can be related to a brightness temperature by using the linear relation of Equation 3.3. While a line can be fitted through any two known points, which makes the calibration process very simple, the determination of the calibration parameters also bears the danger of inconsistencies under rapidly changing measurement conditions, which could lead to large errors in the calculated TBs. The cold chamber measurements described in the previous section are used to investigate the influence of the changing instrument state (due to changing surrounding temperature) on the calibration parameters and the ND signal. To determine a best practice for calibration of MTP raw data, various methods to calibrate HALO-MTP data are being described in the following, giving a brief overview of their respective advantages and disadvantages in connection with the HALO-MTP:

1. Hot-cold calibration, using a cold target (microwave absorber submerged in liquid nitrogen) at temperature T_{cold} and an ambient target (microwave absorber at room temperature) at temperature T_{amb} to derive the calibration parameters. This is the standard calibration method of radiometers in a stable environment. Using this method to calibrate the sensor, before or after taking measurements in the atmosphere, provides the calibration parameters based on two temperatures which lie on the upper edge and below the expected measurement range. Thus, the validity of the calibration for the following measurements can be ensured, as long as the sensor itself is in the same surrounding conditions during the calibration as during the atmospheric measurements, and sufficient instrument stability is given. Since this stability is not given for the MTP instrument, the equations applied for this calibration method are:

$$T_{\mathrm{B}}^{CCh}(c) = s_{\mathrm{cal}}^{CCh}(c_{\mathrm{hot}}) \cdot c - T_{R}^{CCh}(c_{\mathrm{hot}}) \quad (Eq\ 4.1a)$$

$$s_{\mathrm{cal}} = \frac{T_{\mathrm{amb}} - T_{\mathrm{cold}}}{c_{\mathrm{amb}} - c_{\mathrm{cold}}} \quad (Eq\ 4.1b)$$

$$T_R = T_{\mathrm{amb}} - s_{\mathrm{cal}} \cdot c_{\mathrm{amb}} \quad (Eq\ 4.1c)$$

In which c_{hot} denotes a system parameter that describes the instrument state (see following section), so that in-flight data can be related to the Cold-Chamber laboratory measurements within a similar instrument state (indicated by the index “CCh”). Using the hot-cold calibration method is necessary to characterise the noise diode signal used in the second calibration method as described below. However, since it makes use of external calibration targets, the calibration measurement can only be performed on the ground, where single

calibration measurements at arbitrary room temperatures may not be representative of the instrument state during flight, as shown below. However, this method can be used to check the overall health of the instrument in between deployments.

2. MTP built-in hot target (microwave absorber with a heated metal plate in the back) at temperature T_{hot} combined with a noise diode offset signal c_{ND}

$$T_{\mathrm{B}}^{ND}(c) = s_{\mathrm{cal}} \cdot c - (T_{\mathrm{hot}} - s_{\mathrm{cal}} \cdot c_{\mathrm{hot}}) \quad (Eq.\,4.2a)$$

$$s_{\mathrm{cal}} = \frac{T_{\mathrm{ND}}}{c_{\mathrm{ND}} - c_{\mathrm{hot}}} \quad (Eq.\,4.2b)$$

This is the default way to calibrate MTP measurements. By using calibration measurements taken during flight, the calibration roughly follows the individual state of the instrument, whatever conditions the aircraft meets. The down-side of this method is that a faulty noise diode signal can jeopardise reliable calibration. Also, in this method two reference temperatures are used, which are above the expected measurement range: The built-in calibration target is up to 100 K warmer than the outside air temperatures during flight, and T_{ND} is added to this temperature. Hence small uncertainties in the determination of the calibration parameters may lead to large deviations in the calibrated data.

3. MTP built-in hot target combined with HALO static temperature (HALO TS), using the equation

$$T_{\mathrm{B}}^{TS}(c) = s_{\mathrm{cal}} \cdot c - (T_{\mathrm{hot}} - s_{\mathrm{cal}} \cdot c_{\mathrm{hot}}) \quad (Eq.\,4.3a)$$

$$s_{\mathrm{cal}} = \frac{T_{\mathrm{hot}} - TS}{c_{\mathrm{hot}} - c_{0^\circ}} \quad (Eq.\,4.3b)$$

Here c_{0° represents the recorded signal at the horizontal viewing angle, which corresponds to the forward-looking measurement, probing the air masses directly in front of the aircraft. This method is an alternative to the previous calibration method, in the case that the noise diode signal cannot be used. It also follows the individual state of the instrument during measurement flights, but since this method is using the HALO static temperature measurement, the MTP data are no longer independent from the aircraft measurements.

Other calibration methods, such as the tipping curve calibration (e.g. Küchler et al., 2016, or Han and Westwater, 2000) are not available for the DLR MTP ~~due to~~because of the given instrument design~~, and~~ (mainly antenna beam width) and potentially fast-changing atmospheric conditions, due to the moving platform, influencing radiative transfer calculations needed in this approach, as well as the need for an efficient measurement strategy.

In the following, the ~~various~~three methods described above are applied to calibrate MTP data from the cold-chamber measurements (see Section 3.2). First, the hot-cold calibration is used to investigate the temperature-dependence of the calibration parameters themselves (Section 4.1), then, the other calibration methods, which are based on data recorded during mission flights ("in-flight calibration"), are assessed and temperature effects are again discussed (Section 4.2). In section 4.3 the calibration methods are tested and compared to each other using measurements recorded during the ML CIRRUS campaign deployment. A discussion of ~~uncertainties~~measurement uncertainty and summary of the results, leading to an assessment of a best practice for calibration of HALO-MTP data after a campaign, is given is Section 4.4.

4.1 Hot-cold calibration in a cold chamber

When performing cold-target measurements, the interference with a standing wave present between the instrument's receiver hardware and the surface of the slowly evaporating liquid nitrogen (see also Section 4.1.1 in Küchler et al., 2016), was taken into account. As the HALO-MTP is a total-power radiometer (Denning et al., 1989), the output voltage of the detector is proportional to the square of the incoming intensity (Ulaby et al., 1981; Woodhouse, 2005). Thus, the times with least interference of the original signal and the standing waves are defined by minima in the measured signal time series. To find those minima in the cold chamber measurement time series, several steps were taken: (i) a running average (N = 25) is used to minimise the noise on the data; (ii) a spline-fit is used to find a smooth curve, representing the measurements; (iii) the fit is used to interpolate to a higher time-resolution; (iv) the minima of this interpolated curve are used to identify those individual measurement cycles closest to the minima in the time series on which the calibration will be based. Due to noise, the calibration becomes more reliable, if a mean of more than one measurement close to a minimum in the time series is used, hence, the five measurements closest to the time of a minimum in the smooth curve are always included in the analysis.

The resulting calibration parameters are plotted over the corresponding scanning unit temperatures at the time the minimum in the cold target measurements occurred. Fig. 8 clearly shows that the parameters change with the scanning unit temperature. That corroborates the assumption, that HALO-MTP flight data cannot simply be calibrated by using fixed calibration parameters from laboratory measurements at single arbitrary room temperatures, since such measurements are only representative for specific instrument states. Still, it is possible to apply a linear fit to the data, providing a relationship between the MTP scanning unit temperature and the calibration parameters to be used at these temperatures. The same is true when using the hot target measurement signal as a reference, which might better represent the current state of the instrument than the scanning unit temperature. The linear fit parameters are summarised in Table 3.

4.2 Calibration using the MTP built-in target

When applying this (default) calibration method to MTP data, everything builds on the following two assumptions that i) the ND offset signal is the same each time the calibration measurements are performed and ii) the TB measured when pointing towards the heated target corresponds to the measurements of the temperature sensors at the back of the target. Both assumptions are tested in the following, using the calibration measurements performed in the cold-chamber.

Noise diode offset temperature

The ND offset signal is characterised using the hot-cold calibration method used during the cold chamber measurement series, during which the ND is repeatedly activated. Since the calibration parameters are already known from the hot-cold calibration, the temperature offset connected to the signal offset created by the ND can be calculated. Resulting HALO-MTP ND offset temperatures are shown in Fig. 9. The ND offset temperature obviously depends on the count offset resulting from the induced noise on the input signal, which shows a clear dependency on the sensor unit temperature (colouring of the dots in Fig. 9). Again, it is possible to apply a linear fit between the recorded ND offset signal, $\hat{c}_{\mathrm{ND}} = c_{\mathrm{ND}} - c_{\mathrm{hot}}$, and the associated ND offset temperature, derived from the hot-cold calibration method. This fit can be used to find the correct ND offset temperature required in the calibration of mission data. The linear fit values of this correction are shown in Table 4 (last column). In Fig. 9 the deviation of noise diode counts from the linear fit can be seen as being as large as 20 counts (approx. 0.83 K) for any of the three frequency channels. This spread translates into the remaining uncertainty in the ND offset temperature.

Calibration based on outside air temperature

During its deployment in the ML CIRRUS campaign in 2014, occasional failures of the ND, caused by a faulty soldered joint, were experienced. As the ND signal could not be used for calibration, HALO TS can be used instead. This temperature is interpreted as the TB measured at the 0° elevation (horizontal measurement). Simple radiative transfer calculations show, that the MTP measurements at all standard frequency channels are most sensitive to the air directly in front of the sensor (less than 2 km distance; see appendix or Kenntner, 2018, for more details). Thus, the average HALO TS value of the 13 s - period it takes to record an entire MTP measurement cycle (with the 0° measurement being in the middle of the cycle), is representative of the air masses probed by the 0° elevation measurements. Hence, the calibration parameters can also be derived by combining the calibration measurement while pointing at the hot target with the horizontal measurement, using Eq. 4.3.

Hot target temperature measurement

The housekeeping data of the MTP indicate large temperature differences of up to 55 K between the air in front of the hot target and the heated back. Temperature gradients within the absorber material could lead to a misinterpretation of the measured brightness temperature, since the calibration measurement is mostly influenced by the front of the absorber, of which the exact temperature is unknown. There are no temperature sensors built into the absorber material, which could be used to derive the thermal gradient, and measurements with a thermal imager would require disassembling of the instrument, and are no option. Still, to investigate the hot target measurement characteristics, the calibration parameters determined from the hot-cold calibration method, are used to calculate the hot target TB associated with the current measurement signal. Indeed, Figure 8 (bottom panel) shows the clear trend towards colder TBs with lower scanning unit temperatures, which correspond to a colder environment of the MTP instrument, contrary to the readings of two Pt100 temperature sensors, which show the intended target temperature of the heaters placed at the metal back of the target, of just below 45°C during entire mission flights (orange line in Figure 8, bottom panel). The difference can be as large as 3 K. However, the linearity of the sensor again allows for a linear fit between the current scanning unit temperature and the average associated hot target TB. Thus, in-flight calibration can be performed, using a corrected hot target TB, according to the MTP instrument's housekeeping data. The parameters to correct the hot target TBs used in the calibration are shown in Table 5.

4.3 Comparison of calibration methods

There are eight different ways to perform the calibration calculations with and without applying the corrections discussed in the previous sections, summarised in Table 6. All methods are compared to find the best practice of deriving TBs from MTP raw counts by applying all eight methods to the same set of mission data. To do so, segments from all ML CIRRUS mission flights are used, during which the altitude of the aircraft did not change by more than 50 m during a measurement cycle, and no curves were flown (roll smaller than 5°). Note that this definition of usable segments is not based on any parameters connected to the instrument state of the HALO-MTP (e.g. scanning unit temperature), leading to the inclusion of measurement cycles with possibly unstable measurement conditions, e.g. shortly after altitude changes. The only exception is that only those segments are used, during which the ND did not show failures, to ensure comparability of all calibration methods. This way, 38 flight segments with at least 10 minutes duration (i.e. including at least 50 measurement cycles) could be identified. The TBs are calculated based on each individual measurement cycle, but using the calibration coefficients (s_{cal} and T_R) calculated from the average of the relevant data from the seven previous cycles, the seven following cycles, and the cycle itself (N = 15), to account for noise on the calibration measurement signals. As an example, the resulting TBs of the 56.363 GHz measurements at 0° limb-viewing angle during one segment of ML CIRRUS flight MLC10 on April 11th, 2014, are shown in Fig. 10 (top panel). The TBs resulting from all calibration methods show the same time-dependent variations, and mainly differ in

their offset to HALO TS, indicating that differences in the respective calibration coefficients affect the absolute accuracy of the derived TBs more than the precision.

To further investigate the precision of the MTP measurements, a leg-mean value of the HALO TS and the TBs of the 0° elevation measurements is used to determine the offset, which is subtracted from the TBs at all elevation angles:

$$T_{\mathrm{B}}^{corr}(\nu_{LO},\alpha) = T_{\mathrm{B}}(\nu_{\mathrm{LO}},\alpha) - \left(\overline{T_{\mathrm{B}}(\nu_{\mathrm{LO}},0°)} - \overline{TS}\right) \quad (Eq.4.4)$$

with $T_{\mathrm{B}}(\nu_{\mathrm{LO}},\alpha)$ and $T_{B}^{corr}(\nu_{LO},\alpha)$ denoting the original and the corrected TBs under elevation angle α and at a specific frequency channel (i.e. LO), respectively. $\overline{T_{\mathrm{B}}(\nu_{\mathrm{LO}},0°)}$ denotes the leg-mean of the original TBs, measured under 0° elevation, and $\overline{TS}$ represents the leg-mean HALO TS. By using leg-mean values to determine the offset, the corrected TBs will still contain individual small-scale structures, which might differ from those in the HALO TS measurements~~.~~, and, furthermore, large-scale trends of the background atmospheric temperature are also conserved in the resulting data. For individual calibration strategies, the subtracted offset can be as small as 0.8 K or as large as 7 K (see Fig. 10, top panel). The good agreement between all eight corrected TBs under the different viewing angles (see Fig. 10, lower part), indicates that removing the offset will not significantly change the shape of the temperature profile calculated in the retrieval. Moreover, the absolute accuracy of brightness temperature measurements now matches that of the HALO TS, which has an overall ~~uncertainty~~error of 0.5 K (Ungermann et al., 2015).

After applying the offset correction, the RMS difference between the 0° TBs and HALO TS, as shown in Fig. 11, gives a good impression of the capabilities of the different calibration strategies. Naturally, the methods that make use of HALO TS show the smallest deviation from HALO TS readings. However, it is the intention to maintain an independence of the MTP measurements from other measurement systems, increasing the value MTP data adds to the package of instruments flown on HALO. Of those calibration methods that do not use the HALO TS, the most reliable results are obtained when applying the method 'CCH', which uses the calibration values from the hot-cold calibration in the cold chamber measurements, related to the current hot target measurement signal. Whenever reliable ND measurements are available, this calibration method provides equally reliable results. Applying the corrections to T_{ND} or T_{hot} does not significantly change the result, but slightly smaller deviations from HALO TS are seen for the TBs derived using only the T_{hot} correction (method 'TND1b') or both corrections (method 'TND2'). Considering the ND failures during the ML CIRRUS campaign, the favoured calibration strategy is method 'CCH', also applying the offset-correction between the leg-mean 0° TB and the leg-mean HALO TS. The deviation between the resulting 0° elevation TBs and HALO TS is ≤ 0.38 K at all three frequency channels for all ML CIRRUS flight legs with stable instrument conditions. This value is only exceeded when using the calibration method 'CCS' - for all other methods it can be interpreted as the precision of MTP brightness temperature measurements, as will be shown below.

4.4 Discussion of ~~uncertainties and~~HALO-MTP measurement ~~precision~~uncertainty

~~In the literature (e.g. Ulaby et al., 1981; Woodhouse, 2005), the standard formula to derive the measurement uncertainty is defined as the variance of measurement noise, σ_N:~~

~~σ_N~~The overall uncertainty of the MTP measurements includes both systematic errors (e.g. the bias, described above), and random errors (e.g. noise). In case of the HALO MTP, the former can be related to HALO TS, as shown above, the latter mainly influences measurement precision, and influences the ability of the instrument to pick up atmospheric temperature fluctuations. In the literature (e.g. Ulaby et al., 1981), the standard formula to derive radiometric sensitivity (corresponding to measurement precision of an ideal radiometer) is defined as the minimum detectable change in the radiometric antenna temperature of the observed scene:

$$\Delta T_{\text{theo}} = \frac{T_{\text{sys}}}{\sqrt{\Delta f \cdot \tau}} = \frac{T_R + T_{\text{atmo}}}{\sqrt{\Delta f \cdot \tau}} \quad (Eq.\,4.5)$$

in which Δf denotes the filter bandwidth, and τ represents the integration. The HALO-MTP has an ideal filter width of Δf = 200 MHz (see Fig. 2, left panel) and uses an integration time of 200 ms. Assuming a receiver noise temperature of 493.79 K (see Table 7), and a mean atmospheric temperature of 250 K, this leads to a theoretical value of ~~$\sigma_{N,theo}$~~$\Delta T_{\text{theo}} = 0.117\ K$~~,~~. However, these values used to derive the theoretical ~~variance~~value do not take into account, that the effective filter band width is smaller than the ideal value due to small deviations depending on frequency, and because of the gap in the centre of the transmission function (see Fig. 2), so that larger errors are expected for a real measurement system.

In the calibration process, the main uncertainties arise from the use of the different reference temperatures, summarised in Table 7. It is clear that the individual uncertainties assigned to each of the contributing values are not all independent. For example, the uncertainty of the y-intercept (T_R) directly follows from the uncertainty of the slope of the line, but is also directly influenced by the changing instrument state. Hence, a quadratic sum of the individual errors is not suitable and can lead to a large over-estimation of the total TB error. Thus, a sensitivity analysis to estimate the overall measurement uncertainty is performed: reference values (see last column in Table 7) for all parameters with uncertainties are used in a reference calculation. With these values, TBs are calculated for a range of counts between 17500 and 19725, which corresponds to the measurement signal range for atmospheric temperatures, as seen during the ML CIRRUS campaign (approx. 200 K – 300 K). Two control calculations are made, adding the corresponding individual uncertainties (see second-to-last columns in Table 7) in a way that the slope of the calibration line becomes as steep as possible (s_{cal}^{max}, red lines in Fig.12), or as flat as possible (s_{cal}^{min}, blue lines in Fig. 12), following:

$$s_{cal}^{max} = \frac{(T_1 + \Delta T_1) - (T_2 - \Delta T_2)}{(c_1 - c_2) - 2\Delta c} \quad (Eq\ 4.6a)$$

$$T_R^{max} = (T_2 - \Delta T_2) - s_{cal}^{max} \cdot (c_s + \Delta c) \quad (Eq\ 4.6b)$$

$$s_{cal}^{min} = \frac{(T_1 - \Delta T_1) - (T_2 + \Delta T_2)}{(c_1 - c_2) + 2\Delta c} \quad (Eq.4.7a)$$

$$T_R^{min} = (T_2 + \Delta T_2) - s_{cal}^{min} \cdot (c_s - \Delta c) \quad (Eq.4.7b)$$

assuming that T_1 (with associated measurement signal c_1) is the warmer temperature used in the calibration. Comparing the TBs of the reference calculation to those of the two control calculations reveals the maximum uncertainty in the derived TBs.

Furthermore, in parallel to the offset correction introduced in the previous section, a TB correction for the control calculations is introduced: Here, the offset correction is calculated from the difference between the TBs of the control calculation (T_B^{ctr}), and that of the reference calculation (T_B^{ref}), at 18500 cnts (~ 250 K):

$$T_B^{corr}(c) = T_B^{ctr}(c) - \left(T_B^{ctr}(18500) - T_B^{ref}(18500)\right) \quad (Eq.4.8)$$

Results for all three calibration methods are shown in Figure 12. Within the typical region of measurements of one set of elevation angles (vertical, grey shaded region in Fig. 12), the resulting uncertainty is comparable to, or smaller than the expected error from the measurement noise itself ($\Delta T_B = \Delta c \cdot s_{cal}^{ref} - T_{Sys}^{ref}$), indicated by the horizontal black dashed lines. In that, the three approaches to calibrate MTP measurements produce comparable uncertainties in the derived TBs. However, the calibration method relating to the cold chamber measurements is most reliable in the case that the measured signals at certain viewing directions deviate largely from the measurement signal at the horizontal elevation (i.e. if large vertical temperature gradients are present around the current flight level). For all methods, the overall uncertainty is clearly below the already established value of ~ 0.38 K mainly caused by the measurement noise. This is approximately three times larger than the theoretical ~~error~~value, which is expected, as the theory does not consider gain fluctuations (Ulaby et al., 1981). This is not representative of any real radiometric system that is applied outside controlled laboratory conditions, especially the MTP. It also confirms that the uncertainty of derived TBs is dominated by gain fluctuations. Any change of the measurement settings to use larger integration times, and thus reduce this noise, would be at the cost of horizontal resolution.

5 Further considerations relevant for retrieval set-ups

While this work excludes explicit retrieval studies, this section will provide brief insights to a few factors impacting retrieval outputs, which have not yet been mentioned. This includes further sources of input uncertainties as well as some considerations of how different measurement settings would impact the output quality.

Uncertainty from pointing error

The position of the HALO-MTP instrument underneath the wing of the aircraft makes it sensitive to the altitude and speed of the aircraft. The bending and torsion of the aircraft body parts leads to deviations between the assumed pointing of the instrument and its true viewing direction. Adding to this is the aircraft pitch, which also depends on aircraft altitude and speed, introducing a potential source of a systematic error. During flights, the measurement of the inertial sensor, which is part of the HALO-MTP and constantly records the current pitch angle of the instrument, is disturbed by the electromagnetic signal caused by the near-by mounted stepper motor, making the data not reliable enough to allow for a real-time correction of the pointing of the MTP instrument. Thus, the real pointing has to be determined after the flight. Analysing the few reliable data points available after the two campaign deployments revealed that the relative deviation from the true horizontal plane was less than 1° -2° during entire mission flights. Compared to this, the MTP's FOV of 7°-7.5° (see Figure 2 and Section 3.1) is clearly larger. Thus, it is safe to assume that a deviation of the elevation angle of 1° - 2° from the assumed angle does not have a considerable influence on the uncertainty of the retrieval input.

Synthesizer errors

Another error source, which could not be studied in the lab is a possible shift of the LO frequency. To measure the stability of the LO frequency generated by the synthesizer, it would be necessary to disassemble the instrument, which has serious implications for the certification of flying it on the aircraft. However, because of the placement of the LO frequency in the line centre, a small shift will not create serious changes in the recorded TB, however, the measured TB would be caused by a slightly different altitude layer than expected. Because the strongest absorption lines of the absorption band are used, a notable effect would only be seen, if the LO frequency would shift so much that one of the flanks of the line would move out of the filter range~~.~~ (in which case the over-all health of the instrument should be checked). The effect is also dependent on the aircraft altitude, at which the error occurs, due to the effect of line broadening. Furthermore, there is already a smearing effect caused by the FOV of the antenna, which dominates the altitude error, as long as the synthesizer error is small. Further investigation would certainly be useful in the retrieval setup, using the forward radiative transfer modelling.

Measurement settings impacting the retrieval output

There are several settings that can be easily changed when deciding the measurement strategy of the instrument. Both, the set of elevation angles as well as the number and ~~type~~location of frequency channels used in a measurement cycle can be freely chosen. Simple geometric calculations, considering the length of the light path through an altitude layer, lead to a smaller set of elevation angles, which reduces redundancy between measurements due to the field-of-view of the instrument. Details can be found in the appendix.

The frequency of the LO can be set between 48 GHz and 64 GHz (see Table 1). Hence, there are quite a few possibilities to choose different frequency channels, including some at weaker absorption lines than in the standard set-up. Using those, the instrument is able to view deeper into the atmosphere. As a reference, the MTP flown on the ER-2 research aircraft at ~ 20 km altitude only has a measurement range of approximately 2 - 3 km around flight altitude, while using two frequency channels at similarly strong absorption lines to the current instrument (i.e. 57.3 GHz and 58.8 GHz; Gary, 1989). Likewise, the height range of the DC-8 instrument, with LOs at 55.51 GHz, 56.66 GHz, and 58.79 GHz has an 'applicable range' (within which the weighting function drops to 1/e; see appendix) of roughly ±2.8 km (Gary, 2006). In their conclusion of NCAR-MTP data analysis Davis et al. (2014) mention that "it appears that more than about 3 km below the aircraft, the MTP may have difficulty identifying subtle mesoscale variations of temperature".

The choice of frequency channels can have consequences on the calibration options, but can lead to a clearly enhanced ~~vertical observation~~altitude range of sensitivity of the MTP. A pencil-beam calculation of the weighting functions at different frequency channels and viewing angles shows that changing the measurement strategy to include four frequency channels (instead of three) and only eight viewing angles (instead of ten) can significantly increase the vertical range and resolution of MTP measurements while still maintaining the length of one measurement cycle (i.e. not changing the horizontal resolution). However, an in-depth assessment of alternative measurement strategies must include forward radiative transfer modelling, using the information given in Section 3, and would go beyond the scope of this study. Details of the pencil-beam radiative transfer calculations and some implications for the calibration strategy are shown in the appendix.

6 Summary

This ~~work~~study shows a thorough investigation of the MTP instrument operated by DLR and flown on the HALO aircraft. It is the first time a thorough characterisation of a MTP instrument, including the assessment of brightness temperature calibration, is published. The knowledge of the instrument characteristics, such as the instrument transmission function and the antenna diagram, its calibration and related measurement uncertainties, as well as the vertical observation range, is fundamental to the correct set-up of a retrieval algorithm, as well as for the interpretation of the retrieved temperature signals and conclusions about the atmospheric conditions around flight altitude.

Using the standard measurement settings, the instrument response function was determined along with the antenna diagram (Section 3). The results show symmetric shapes of all transmission functions, which is the desired and expected result. While small side-lobes are detected in the antenna diagram, the main lobe has a symmetrical Gaussian shape, with a full-width-half-maximum of 7.0° - 7.5°, which represents the field of view of the

instrument. A smaller field of view could only be achieved by using a different shape of the rotating mirror as well as a larger horn antenna, which the compact design, needed for the wing-carrier instrument, does not allow for.

Measurements in a cold-chamber set-up, as well as data recoded during a field campaign, were used to characterise the MTP's noise figurecharacteristics (Section 3). It can be described by a Gaussian distribution with mean of zero counts, and a standard deviation of 6 counts. (slightly less than 0.25 K). It was shown that the measurement noise can be characterised as a red noise, with lag-1 auto-correlation of 0.7, which indicates that a time-series of MTP data may show wave-like structures caused by internal noise.; however, the presented characterisation of the HALO-MTP noise allows the identification of significant atmospheric signals in MTP measurement time series, as long as the amplitude of the atmospheric signal is larger than the measurement precision (i.e. significantly larger than 0.38 K), and wave-like signals (e.g. caused by gravity waves) can be clearly separated from the noise-induced structures, caused by the auto-correlation of MTP measurements.

Furthermore, in the laboratory measurements the linear relationship between the instrument's measurements and the source temperature could be confirmed. Based on this linear relationship, the calibration of MTP raw data to derive brightness temperatures is possible, and was further analysed (Section 4). The measurements revealed clear changes in all calibration parameters, depending on the cold chamber temperature. This includes a change in the measured brightness temperature when pointing towards the built-in hot calibration target, as well as a change in receiver noise temperature caused by the electrical parts, and the calibration slope, caused by a change in amplification of the signal. Corrections to account for those changes have been found, and it could be shown, that with the application of those corrections, brightness temperatures could be derived from ML CIRRUS mission data, with an accuracy matching that of the HALO static temperature, and a precision better than 0.38 K. To achieve this accuracy, theThe necessity of an offset-correction relative to HALO TS has been identified, both for the calibration relating to laboratory measurements, and using the noise diode signal. The correction procedure was introduced as comparing the leg-mean of the calculated TBs at 0° elevation angle (horizontal measurement) to the leg-mean of HALO TS. While removing a potential bias on the measured brightness temperatures, this method conserves trends in the background atmospheric temperature, as frequently observed during longer flight legs.

It was shown that all presented calibration methods produce comparable results. Considering the desire for MTP measurements mostly independent from other measurements (such as the HALO TS, which can then be used as reference), and technical problems with the ND, experienced during the ML CIRRUS campaign in 2014, the favoured method of calibration is to use calibration parameters from the cold chamber measurement series, linked to the system state via the measurement signal while pointing towards the MTP built-in target. When analysing the uncertainty of the calibrated brightness temperatures (Section 4.4), it was found that this method performs best, whenever large vertical temperature gradients are

present near flight level. Furthermore, the analysis of uncertainties of the calibration parameters shows that it is clearly dominated by the contribution from measurement noise. Other uncertainties, such as the pointing of the instrument, or synthesizer errors are negligible compared to this uncertainty.

A brief discussion has been given on possibilities to further improve the quality and value of MTP measurements (Section. 5), by changing the measurement set-up within the given possibilities allowed by the instrument hardware. Simple estimations indicate that the signal is mostly influenced by the first 1.5 - 2 km distance to the aircraft altitude, both above and below flight level. The instrument hardly collects any usable information on the state of the atmosphere outside of the resulting ~ 3 km region around flight altitude (i.e. ±1.5 km around flight level). A proposal to improve the measurement strategy for future missions of the MTP is made, involving a reduction of the number of elevation angles used and including frequencies of weaker absorption lines of the 60 GHz oxygen absorption complex. The considerations shown in the appendix indicate that the range of sensitivity above the aircraft can be increased to at least 2 km, and up to approximately 4-5 km below the aircraft at an aircraft altitude of 11 km. At the same time the horizontal resolution of MTP measurements can be maintained. This is a significant improvement in the value of MTP data.

Overall, this ~~work shows all necessary~~study summarises the investigation of instrument parameters and characteristics ~~needed~~, necessary to accurately analyse and interpret the data produced by HALO-MTP measurements. It is the basis to understand measurement uncertainty, the (vertical) range in which derived atmospheric properties are valid, to identify significant atmospheric signals in times-series of HALO-MTP data, and a guideline for choosing the best-possible strategy to record and calibrate mission data. Using this information, the best-possible data input for the retrieval algorithm, used to derive absolute temperature profiles, can be obtained. With that basis the HALO-MTP can provide valuable information on the atmospheric state which can be utilised in many studies on atmospheric dynamics or in connection to in-situ as well as other remote-sensing measurements made on the same mission flights.

Data availability

The study is based on the analysis of raw measurement data from the HALO-MTP, which is available from the first author upon request. HALO-TS data is available at the HALO database at https://halo-db.pa.op.dlr.de/mission/2.

Author contribution

M. Heckl prepared the manuscript and figures with contributions from all coauthors. Moreover, M. Heckl has recorded the MTP data, has written the code required to interpret the measurements and determine the best calibration methods. A. Fix and M. Rapp contributed to the interpretation of the measurements and determination of results. M. Jirousek was involved with the laboratory measurements and overall instrument set-up. F. Schreier and J. Xu have prepared the radiative transfer code used in the Appendix, and were involved in the works related to radiative transfer studies.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

This work was partly supported by the Bundesministerium für Bildung und Forschung (BMBF) under project 01LG1206C (ROMIC/GW-LCYCLE). Fruitful discussions with Manfred Birk (DLR-MF), Martin Hagen (DLR-IPA), and Harald Czekala (RPG) have also contributed to the work presented in this study. Thanks to MJ Mahoney, Dick Denning, and Boon Lim (NSA-JPL) and to Julie Haggerty (NCAR) for helpful discussions and support in MTP operation.

References

Anderson, G.; Clough, S.; Kneizys, F.; Chetwynd, J.; Shettle, E. AFGL Atmospheric Constituent Profiles (0–120 km); Technical Report TR-86-0110; AFGL: Hanscom AFB, MA, USA, 1986

Bacmeister, J.T, Schoeberl, M. R., Lait, L., Newman, P., and Gary, B.: Small-Scale Waves Encountered During AASE, Geophysical Research Letters, 17:349-352, 1990.

Bacmeister, J.T., Eckermann, S., Newman, P., Lait, L., Chan, K., Loewenstein, M., Proffit, M., and Gary, B.: Stratospheric horizontal wavenumber spectra of winds, potential temperature, and atmospheric tracers observed by high-altitude aircraft, Journal of Geophysical Research, 101:9441-9470, 1996.

Bacmeister, J.T., Eckermann, S. D., Tsias, A., Carslaw, K. S., and Peter, T.: Mesoscale Temperature Fluctuations Induced by a Spectrum of Gravity Waves: A Comparison of Parameterizations and Their Impact on Stratospheric Microphysics, Journal of the Atmospheric Sciences, 56:1913-1924., 1999.

Balanis, C.A.: Antenna theory-: analysis and design, 2. ed., New York, NY : Wiley, 1997.

Chan, K., Pfister, L., Bui, T., Bowen, S., Dean-Day, J., Gary, B., Fahey, D., Kelly, K., Webster, C., and May, R.: A Case Study of the Mountain Lee Wave Event of January 6, 1992. Geophysical Research Letters, 20:2551-2554, 1993.

Cho, J., Newell, R., Bui, T., E.V. Browel and, M. F., Mahoney, M., G.L.Gregory, Sachse, G., S.A.Vay, Kucsera, T., and Thompson, A.M.: Observations of convective and dynamical instabilities in tropopause folds and their contribution to stratosphere-troposphere exchange. Journal of Geophysical Research, 104:21,549-21,568, 1999.

Corti, T., Luo, B., de Reus, M., Brunner, D., Cairo, F., Mahoney, M., Martucci, G., Matthey, R., Mitev, V., dos Santos, F., Schiller, C., Shur, G., Sitnikov, N., Spelten, N., V¨ossing, H., Borrmann, S., and Peter, T.: Unprecedented evidence for deep convection hydrating the tropical stratosphere. Geophysical Research Letters, 35(10):L10810, 2008.

Davis, C.A., Ahijevych, D.A., Haggerty, J A., and Mahoney, M.J.: Observations of Temperature in the Upper Troposphere and Lower Stratosphere of Tropical Weather Disturbances. Journal of the Atmospheric Sciences, 71(5):1593-1608, 2014.

Dean-Day, J., Chan, K., Bowen, S., Bui, T., Gary, B., and Mahoney, M.: Dynamics of Rocky Mountain lee waves observed during SUCCESS. Geophysical Research Letters, 25:1351-1354, 1998.

Denning, R.F., Guidero, S.L., Parks, G.S., and Gary, B.L.: Instrument Description of the Airborne Microwave Temperature Profiler. Journal of Geophysical Research, 94(D14):16,757-16,765, 1989.

Dörnbrack, A., Birner, T., Fix, A., Flentje, H., Meister, A., Schmid, H., Browell, E.V., and Mahoney, M.J.: Evidence for inertia gravity waves forming polar stratospheric clouds over Scandinavia. Journal of Geophysical Research, 107:8287, 2002.

Eckermann, S., Dörnbrack, A., Flentje, H., Vosper, S.B., Mahoney, M.J., Bui, T.P., and Carslaw, K.S.: Mountain Wave-Induced Polar Stratospheric Cloud Forecasts for Aircraft Science Flights during SOLVE/THESEO 2000. Weather and Forecasting, 21:42-68, 2006.

Gamblin, B., Toon, O., Tolbert, M., Kondo, Y., Takegawa, N., Irie, H., Koike, M., Ballenthin, J., Hunton, D., Miller, T., Viggiano, A., Anderson, B., Avery, M., Sachse, G., Podolske, J., Guenther, K., Sorenson, C., and Mahoney, M.: Nitric acid condensation on ice: 1. Non-HNO3 constituent of NOY condensing cirrus particles on upper tropospheric. Journal of Geophysical Research, 111:D21203, 2006.

Gary, B. L.: Observational Results Using the Microwave Temperature Profiler During the Airborne Antarctic Ozone Experiment. Journal of Geophysical Research, 94(D9):11,223-11,231, 1989.

Gary, B.: Mesoscale temperature fluctuations in the stratosphere. Atmospheric Chemistry and Physics, 6:4577-4589, 2006.

Gary, B.: Mesoscale temperature fluctuations in the Southern Hemisphere stratosphere. Atmospheric Chemistry and Physics, 8:4677-4681, 2008.

Haggerty, J., Schick, K., Mahoney, M. J., and Lim, B.: The NCAR Microwave Temperature Profiler: Data applications from recent deployments. In 2014 13th Specialist Meeting on Microwave Radiometry and Remote Sensing of the Environment (MicroRad), pages 133-135, 2014.

Han, Y. and Westwater, E.R.: Analysis and improvement of tipping calibration for ground-based microwave radiometers. IEEE Trans. Geosci. Remote. Sens. 38(3): 1260-1276, 2000.

Hartmann, D., Chan, K., Gary, B., Schoeberl, M., Newman, P., Martin, R., Lowenstein, M., Podolske, J., and Strahan, S.: Potential Vorticity and Mixing in the South Polar Vortex During Spring. Journal of Geophysical Research, 94(D9):11,223-11,231, 1989.

Jacob, M., Ament, F., Gutleben, M., Konow, H., Mech, M., Wirth, M., and Crewell, S.: Investigating the liquid water path over the tropical Atlantic with synergistic airborne measurements. Atmos. Meas. Tech., 12, 3237–3254, 2019.

Jensen, E., Pfister, L., Bui, T.-P., Lawson, P., and Baumgardner, D~~.~~.: Ice nucleation and cloud microphysical properties in tropical tropopause layer cirrus. Atmospheric Chemistry and Physics, 10(3):1369-1384, 2010.

Kenntner, Mareike ~~(2018):~~ Using MTP measurements to characterise atmospheric gravity waves in the Tropopause region. Dissertation, DLR, Institut für Physik der Atmosphäre and LMU München, 2018.

Krautstrunk M., and Giez A.: The Transition From FALCON to HALO Era Airborne Atmospheric Research. In: Schumann U. (eds) Atmospheric Physics. Research Topics in Aerospace. Springer, Berlin, Heidelberg, 2012.

Küchler, N., ~~D. D.~~ Turner, ~~U.~~D.D., Löhnert, U., and ~~S.~~ Crewell ~~(2016),~~S.: Calibrating ground-based microwave radiometers: Uncertainty and drifts, Radio Sci., 51, 311–327, ~~doi:10.1002/2015RS005826.~~2016.

Leutbecher, M. and ~~Volkert~~Volkert H.: The Propagation of Mountain Waves into the Stratosphere: Quantitative Evaluation of Three-Dimensional Simulations. Journal of the Atmospheric Sciences, 57:3090-3108, 2000.

Liebe, H., Rosenkranz, P., and Hufford, G.: Atmospheric 60-GHz oxygen spectrum: New laboratory measurements and line parameters. Journal of Quantitative Spectroscopy and Radiative Transfer, 48(5):629 – 643, 1992.

Lim, B., Mahoney, M., Haggerty, J., and Denning, R.: The Microwave Temperature Profiler performance in recent airborne campaigns. In 2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS, pages 3363-3366, 2013.

Mahoney, M. and Denning, R.: A State-of-the-Art Airborne Microwave Temperature Profiler (MTP~~) .~~). In 33rd International Symposium on the Remote Sensing of the Environment, 2009.

Marcy, T., Popp, P., Gao, R., Fahey, D., Ray, E., Richard, E., Thompson, T., Atlas, E., Loewenstein, M., Wofsy, S., Park, S., Weinstock, E., Swartz, W., and Mahoney, M.: Measurements of trace gases in the tropical tropopause layer. Atmospheric Environment, 41(34):7253 – 7261, 2007.

McGrath, A. and T. Hewison: Measuring the Accuracy of MARSS—An Airborne Microwave Radiometer. J. Atmos. Oceanic Technol., 18, 2003–2012, 2001

Mech, M., Orlandi, E., Crewell, S., Ament, F., Hirsch, L., Hagen, M., Peters, G., and Stevens, B.: HAMP – the microwave package on the High Altitude and LOng range research aircraft (HALO), Atmos. Meas. Tech., 7, 4539–4553, https://doi.org/10.5194/amt-7-4539-2014, 2014.

Murphy, D. and Gary, B.: Mesoscale Temperature Fluctuations and Polar Stratospheric Clouds. Journal of Atmospheric Sciences, 52:1753 – 1760, 1995.

Nielsen-Gammon, J., Powell, C., Mahoney, M., Angevine, W., Senff, C., White, A., Berkowitz, C., Doran, C., and Knupp, K.: Multisensor Estimation of Mixing Heights over a Coastal City. Journal of Applied Meteorology and Climatology, 47:27 – 43, 2008.

Pfister, L., Chan, K., Bui, T., Bowen, S., Legg, M., Gary, B., Kelly, K., Proffitt, M., and Starr, W.: Gravity Waves Generated by a Tropical Cyclone During the STEP Tropical Field Program: A Case Study. Journal of Geophysical Research, 98:8611 – 8638, 1993.

Popp, P., Marcy, T., Jensen, E., Karcher, B., Fahey, D., Gao, R., Thompson, T., Rosenlof, K., Richard, E., Herman, R., Weinstock, E., Smith, J., May, R., Wilson, J., Heymsfield, A., Mahoney, M., and Thompson, A.: The observation of nitric-acid containing particles in the tropical lower stratosphere. Atmospheric Chemistry and Physics, 6:601 – 611, 2006.

Rothman, L., Rinsland, C., Goldman, A., Massie, S., Edwards, D., Flaud, J.-M., Perrin, A., Camy-Peyret, C., Dana, V., Mandin, J.-Y., Schroeder, J., McCann, A., Gamache, R., Wattson, R., Yoshino, K., Chance, K., Jucks, K., Brown, L., Nentchinov, V., and Varanas, P. (1998). ~~THE~~The HITRAN ~~MOLECULAR SPECTROSCOPIC DATABASE AND~~Molecular Spectroscopic Database and HAWKS (~~HITRAN ATMOSPHERIC WORKSTATION~~Hitran Atmospheric WorKStation): 1996 ~~EDITION~~Edition. J. Quant. Spectrosc. Rad. Transfer, 60(5):665-710.

Schreier, F.; Gimeno García, S., Hochstaffl, P., Städt, S~~.~~.: Py4CAtS—PYthon for Computational ATmospheric Spectroscopy. Atmosphere ~~2019~~, 10, 262. 2019.

Schreier, F., Gimeno García, S., Hedelt, P., Hess, M., Mendrok, J., Vasquez, M., and Xu, J.: GARLIC a general purpose atmospheric radiative transfer line-by-line infrared-microwave code: Implementation and evaluation. Journal of Quantitative Spectroscopy and Radiative Transfer, 137:29 – 50, 2014.

Schumann, U., Kiemle, C., Schlager, H., Weigel, R., Borrmann, S., D'Amato, F., Krämer, M., Matthey, R., Protat, A., Voigt, C., and Volk, C.: Long-lived contrails and convective cirrus above the tropical tropopause. Atmospheric Chemistry and Physics, 17(3):2311-2346, 2017.

Schwarz, J., Spackman, J., Fahey, D., Gao, R., Lohmann, U., Stier, P., Watts, L., Thomson, D., Lack, D., Pfister, L., Mahoney, M., Baumgardner, D., Wilson, J., and Reeves, J.: Coatings and their enhancement of black carbon light absorption in the tropical atmosphere. Journal of Geophysical Research: Atmospheres, 113(D3):D03203, 2008.

Sitnikova, V., Sitnikov, N., Ulanovskii, A., Shur, G., Lukyanov, A., and F. Ravegnani, M.M.: Estimation of the Tropospheric Air Ratio near the Thermal Tropopause Using the Aircraft Measurements. Russian Meteorology and Hydrology, 34(8):510 – 514, 2009.

Spinei, E., Cede, A., Herman, J., Mount, G., Eloranta, E., Morley, B., Baidar, S., Dix, B., Ortega, I., Koenig, T., and Volkamer, R.: Ground-based direct-sun DOAS and airborne MAX-DOAS measurements of the collision-induced oxygen complex, O2O2, absorption with significant pressure and temperature differences. Atmospheric Measurement Techniques, 8(2):793-809, 2015.

Tabazadeh, A., Toon, O., Gary, B., Bacmeister, J., and Schoeberl, M.: Observational constraints on the formation of type ia polar stratospheric clouds. Geophysical Research Letters, 23(16):2109 – 2112, 1996.

Thornton, B.F., Toohey, D.W., Tuck, A.F., Elkins, J.W., Kelly, K.K., Hovde, S.J., Richard, E.C., Rosenlof, K.H., Thompson, T.L., Mahoney, M.J., and Wilson, J.C.: Chlorine activation near the midlatitude tropopause. Journal of Geophysical Research: Atmospheres, 112(D18):D18306, 2007.

Torrence, C. and Compo, G.P.: A Practical Guide to Wavelet Analysis. Bulletin of the American Meteorological Society, 79(1):61 – 78, 1998.

Tuck, A., Baumgardner, D., Chan, K., Dye, J., Elkins, J., Hovde, S., Kelly, K., Loewenstein, M., Margitan, J., May, R., Podolske, J., Proffitt, M., Rosenlof, K., Smith, W., Webste, C., and Wilson, J.: The Brewer-Dobson Circulation In the Light of High Altitude In Situ Aircraft Observations. Quarterly Journal of the Royal Meteorological Society, 123:1 – 69, 1997.

Tuck, A., Hovde, S., Kelly, K., Mahoney, M., Proffitt, M., Richard, E., and Thompson, T.: Exchange between the upper tropical troposphere and the lower stratosphere studied with aircraft observations. Journal of Geophysical Research, 108:4734, 2003.

Ungermann, J., Blank, J., Dick, M., Ebersoldt, A., Friedl-Vallon, F., Giez, A., Guggenmoser, T., ~~H¨opfner~~Höpfner, M., Jurkat, T., Kaufmann, M., Kaufmann, S., Kleinert, A., ~~Kr¨amer~~Krämer, M., Latzko, T., Oelhaf, H., Olchewski, F., Preusse, P., Rolf, C., Schillings, J., SuminskaEbersoldt, O., Tan, V., Thomas, N., Voigt, C., Zahn, A., Zöger, M., and Riese, M.: Level 2 processing for the imaging Fourier transform spectrometer GLORIA derivation and validation of temperature and trace gas volume mixing ratios from calibrated dynamics mode spectra. Atmospheric Measurement Techniques, 8(6):2473-2489, 2015.

Ulaby, F., Moore, R., and Fung, A.: Microwave remote sensing: active and passive. Volume I: microwave remote sensing fundamentals and radiometry. Addison-Wesley; Remote Sensing Series 2, 1981.

Urbanek, B., Groß, S., Schäfler, A., and Wirth, M.: Determining stages of cirrus evolution: a cloud classification scheme. Atmospheric Measurement Techniques, 10(5):1653 – 1664, 2017.

Voigt, C., Schumann, U., Minikin, A., Abdelmonem, A., Afchine, A., Borrmann, S., Boettcher, M., Buchholz, B., Bugliaro, L., Costa, A., Curtius, J., Dollner, M., Dörnbrack, A., Dreiling, V., Ebert, V., Ehrlich, A., Fix, A., Forster, L., Frank, F., Fütterer, D., Giez, A., Graf, K., Grooß, J.-U., Groß, S., Heimerl, K., Heinold, B., Hüneke, T., Järvinen, E., Jurkat, T., Kaufmann, S., Kenntner, M., Klingebiel, M., Klimach, T., Kohl, R., Krämer, M., Krisna, T. C., Luebke, A., Mayer, B., Mertes, S., Molleker, S., Petzold, A., Pfeilsticker, K., Port, M., Rapp, M., Reutter, P., Rolf, C., Rose, D., Sauer, D., Schäfler, A., Schlage, R., Schnaiter, M., Schneider, J., Spelten, N., Spichtinger, P., Stock, P., Walser, A., Weigel, R., Weinzierl, B., Wendisch, M., Werner, F., Wernli, H., Wirth, M., Zahn, A., Ziereis, H., and Zöger, M.: ML-CIRRUS: The Airborne Experiment on Natural Cirrus and Contrail Cirrus with the High-Altitude Long-Range Research Aircraft HALO . Bulletin of the American Meteorological Society, 98(2):271 – 288, 2017.

Wang, L., Alexander, M., Bui, T., and Mahoney, M.: Small-scale gravity waves in ER-2 MMS/MTP wind and temperature measurements during CRYSTAL-FACE. Atmospheric Chemistry and Physics, 6:1091 – 1104, 2006.

Wirth, M., Fix, A., Mahnke, P. *et al.* The airborne multi-wavelength water vapor differential absorption lidar WALES: system design and performance. *Appl. Phys. B* **96,** 201, 2009.

Woodhouse, I.H. Introduction to Microwave Remote Sensing. CRC press, 2005.

Xu, J., Schreier, F., Doicu, A., and Trautmann, T.: Assessment of Tikhonov-type regularization methods for solving atmospheric inverse problems. Journal of Quantitative Spectroscopy and Radiative Transfer, 184:274 – 286, 2016.

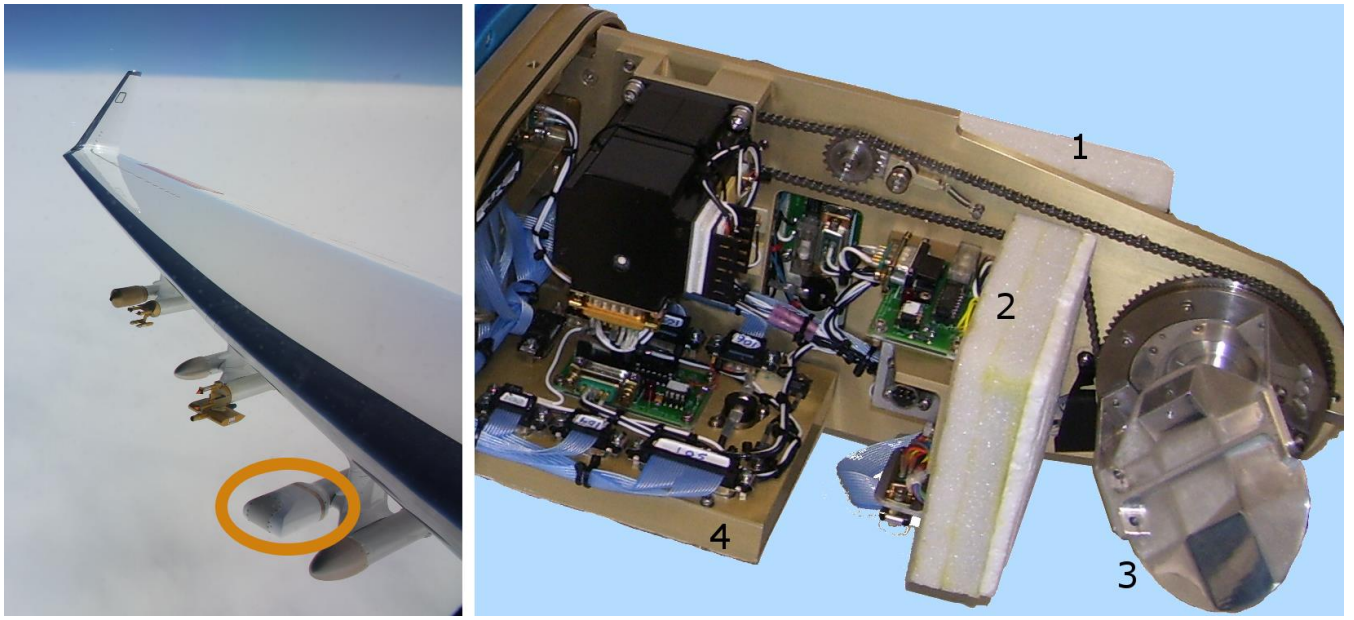


Figure 1: HALO-MTP instrument. Left: Position of the MTP underneath the wing of the HALO aircraft. Right: MTP sensor unit in the lab. Marked with numbers are the radiometer unit (1), the hot calibration target (2), the rotating mirror (3) and the electronic unit (4), which contains various temperature sensors such as the scanning unit temperature or the pod air temperature sensor.

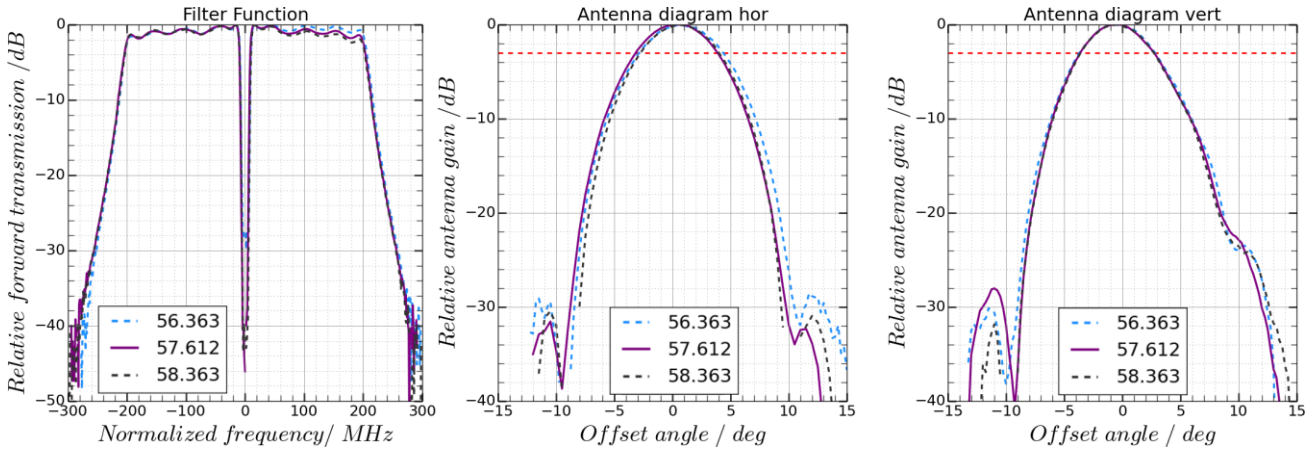


Figure 2: HALO-MTP Filter functions (left panel) and antenna diagram of the horizontal (middle) and vertical (right) plane recorded at standard measurement frequencies. Red dashed lines indicate the half maximum value. All data are normalised so that the maximum value shown is 0.

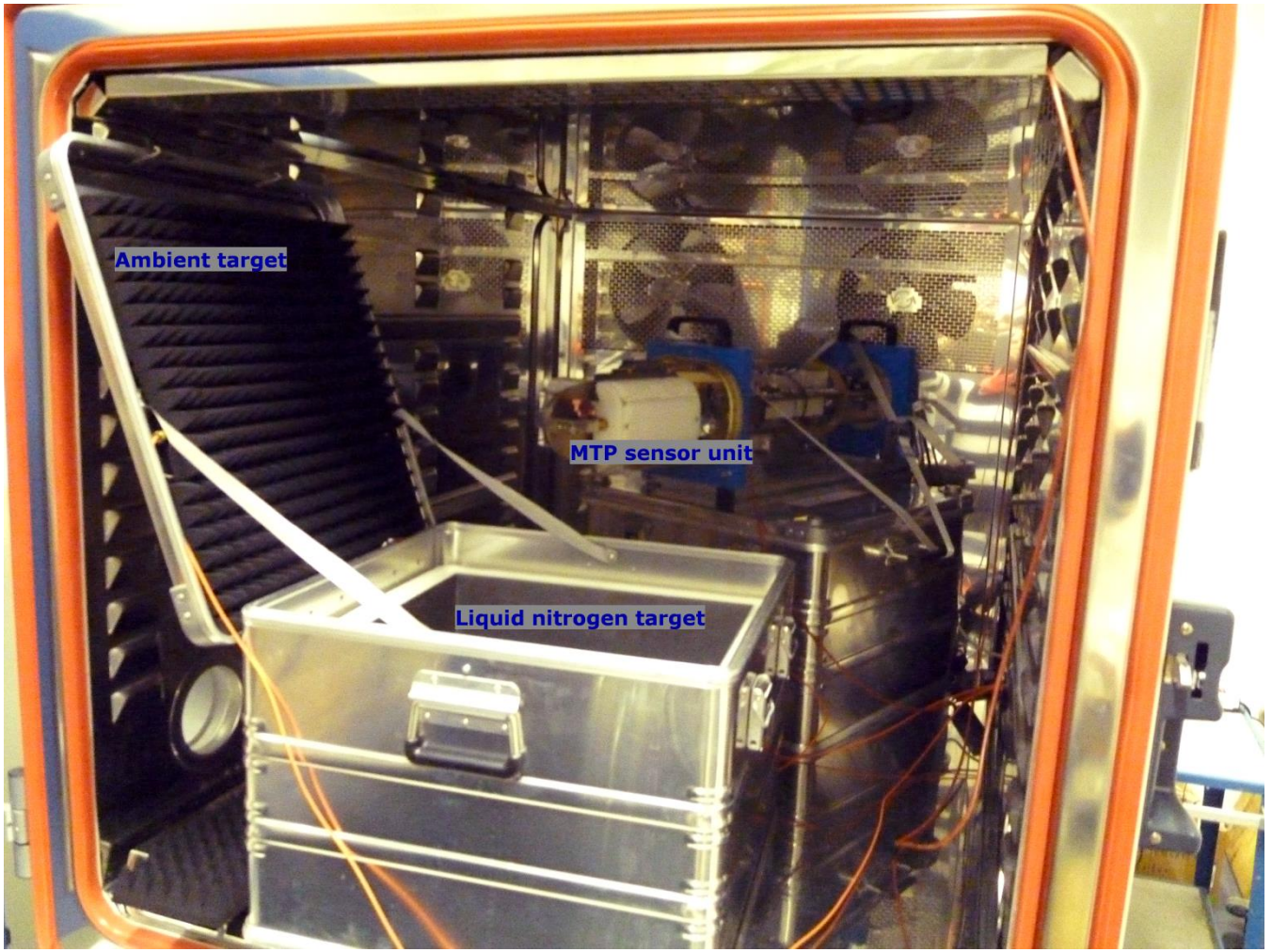


Figure 3: HALO-MTP inside the cold-chamber. For the measurements, the box containing the liquid nitrogen and the ambient target was rotated to face towards the MTP sensor unit. A second microwave absorber was placed on the ceiling of the chamber to function as second ambient target.

5

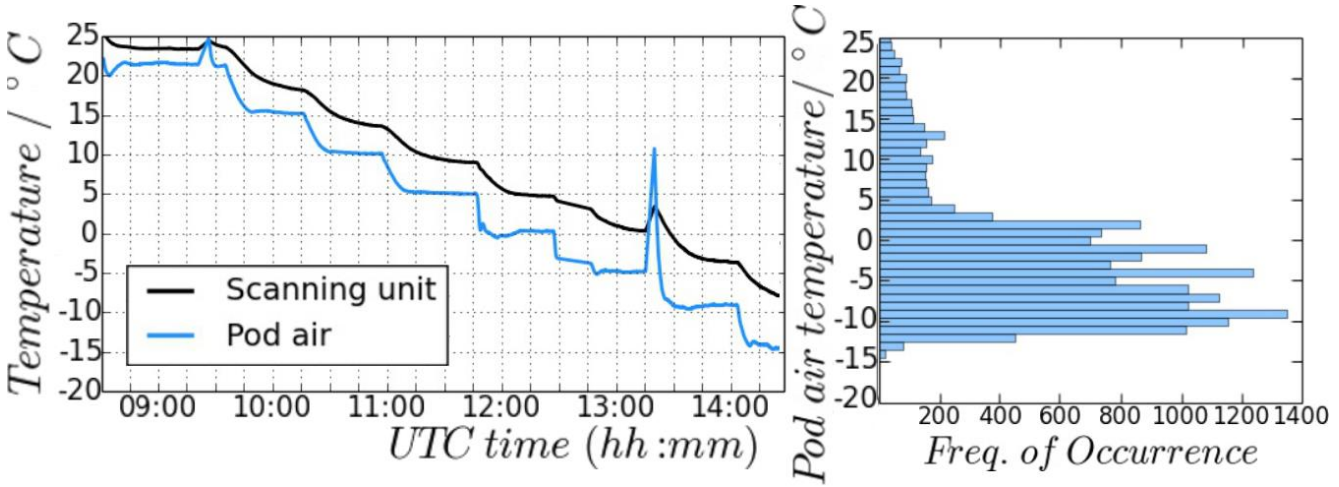


Figure 4: Temperature sensor measurements during cold chamber measurements (left panel; black line: Scanning unit sensor, blue line: Pod air sensor). At 0 °C (~11:45 UTC) the cold chamber software had to be restarted, causing a longer stabilisation period, and at -5 °C (~13:15 UTC) the cold chamber was opened to re-fill the liquid nitrogen in the cold target causing the spikes in the temperature measurement. Right panel: Pod air temperature measurements during all ML CIRRUS campaign flights.

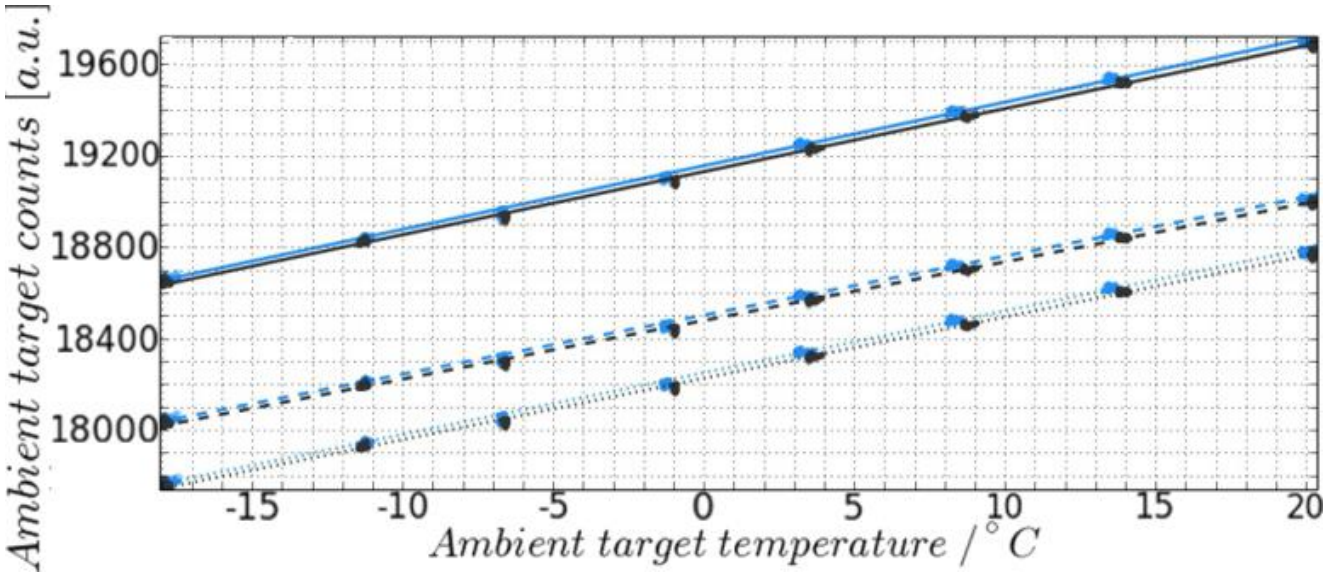


Figure 5: Ambient target temperature vs. measured signal (counts) of the two ambient targets for all three standard frequency ~~channelsof~~channels of the HALO-MTP (different line styles). Different line colours correspond to the measurement of the two individual ambient targets.

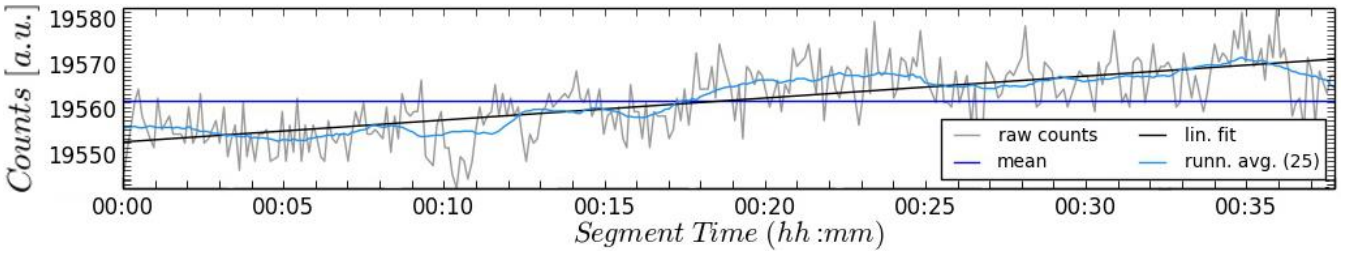


Figure 6: Measured signal (grey line) at 56.363 GHz while looking at the ambient target inside the cold chamber as well as a running average (light, blue line), mean value (blue line) and linear fit (black line). The corresponding brightness temperature change in the linear fit during the segment is about 0.8 K.

5

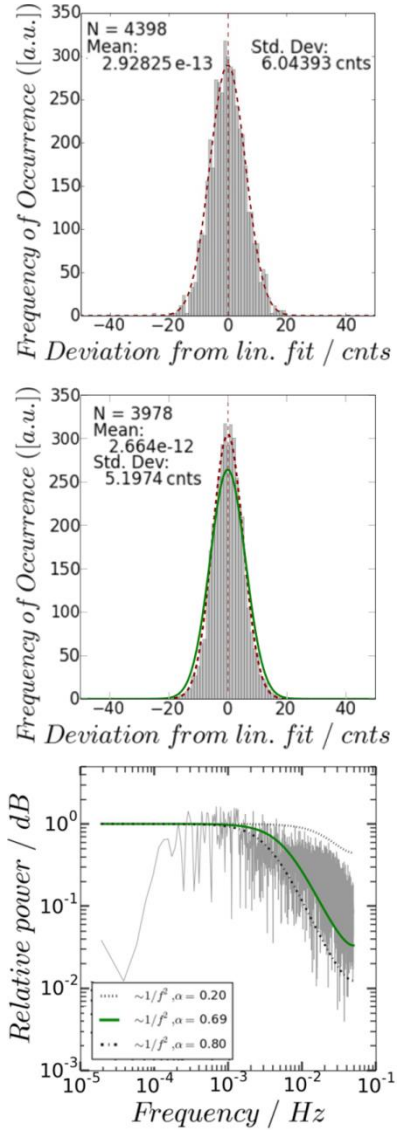


Figure 7: HALO-MTP noise ~~figure~~characteristics at 56.363 GHz. Figures of the other two standard frequency channels look similar. Fit parameters for all three frequency channels are summarised in Table 2. ~~Red dashed line: Gauss-fit to data.~~ Top: Laboratory measurements in cold chamber. Red dashed line: Gauss-fit to data. Middle: derived from ML CIRRUS flight data. Green line: Ideal Gauss function with the mean at 0.0 cnts and 6 cnts (~0.25 K) standard deviation, as implied by the cold chamber noise ~~figure~~spectrum. Bottom: Noise spectrum calculated from ML CIRRUS flight data. Black, dashed

lines: theoretical power spectra of $1/f^2$ noise with lag1-correlations of $\alpha = 0.2$ and $\alpha = 0.8$. Green, solid line: theoretical power spectrum of $1/f^2$ noise with lag1-correlation of input data.

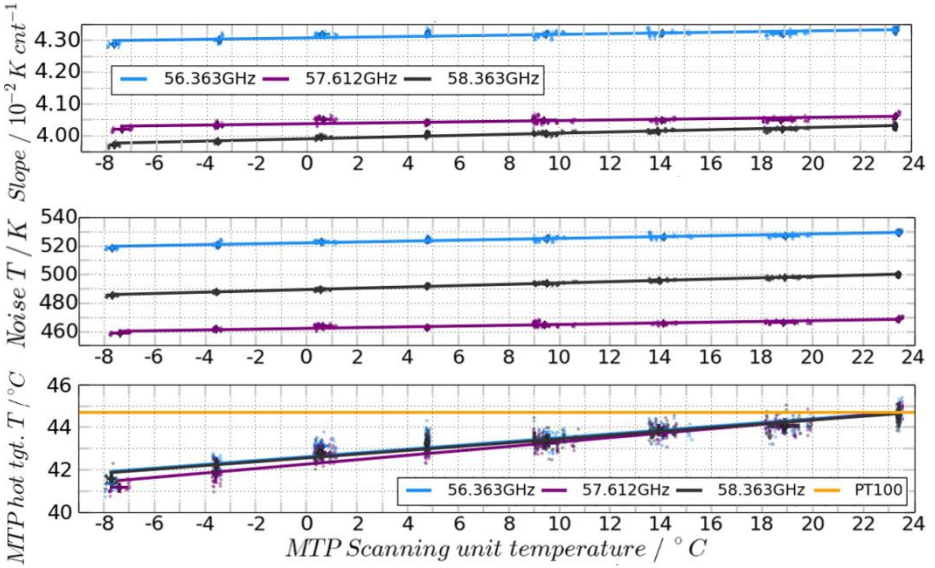


Figure 8: Calibration parameters calculated from hot-cold calibration for standard frequency channels during cold chamber measurements. Top: slope of calibration line. Middle: Receiver noise temperature (T_R). Bottom: Calculated hot target TBs at different scanning unit temperatures during cold chamber measurements. Small dots: Single measurements contributing to the average at one scanning unit temperature. Orange line: Pt100 sensor readings.

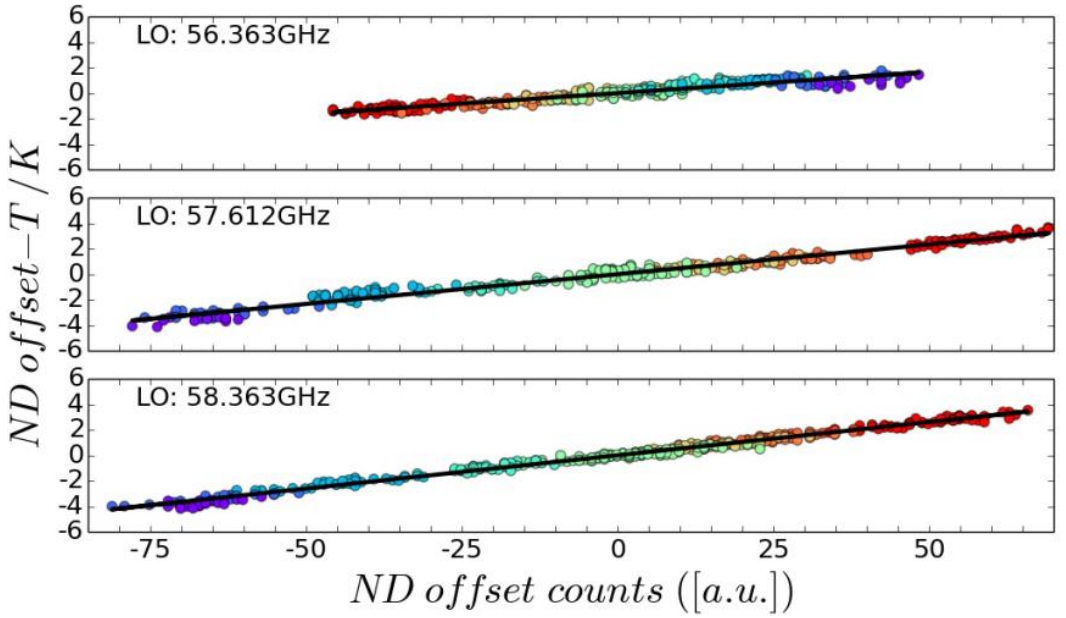


Figure 9: Calculated ND offset temperature for occurring ND count offsets (For better comparability, the means of the temperature and count values have been removed – mean values are shown in Table 4). Colour-coding: MTP scanning unit temperature (Blue: colder - red: warmer). Black line: Linear fit, linking ND offset temperature to offset counts.

5

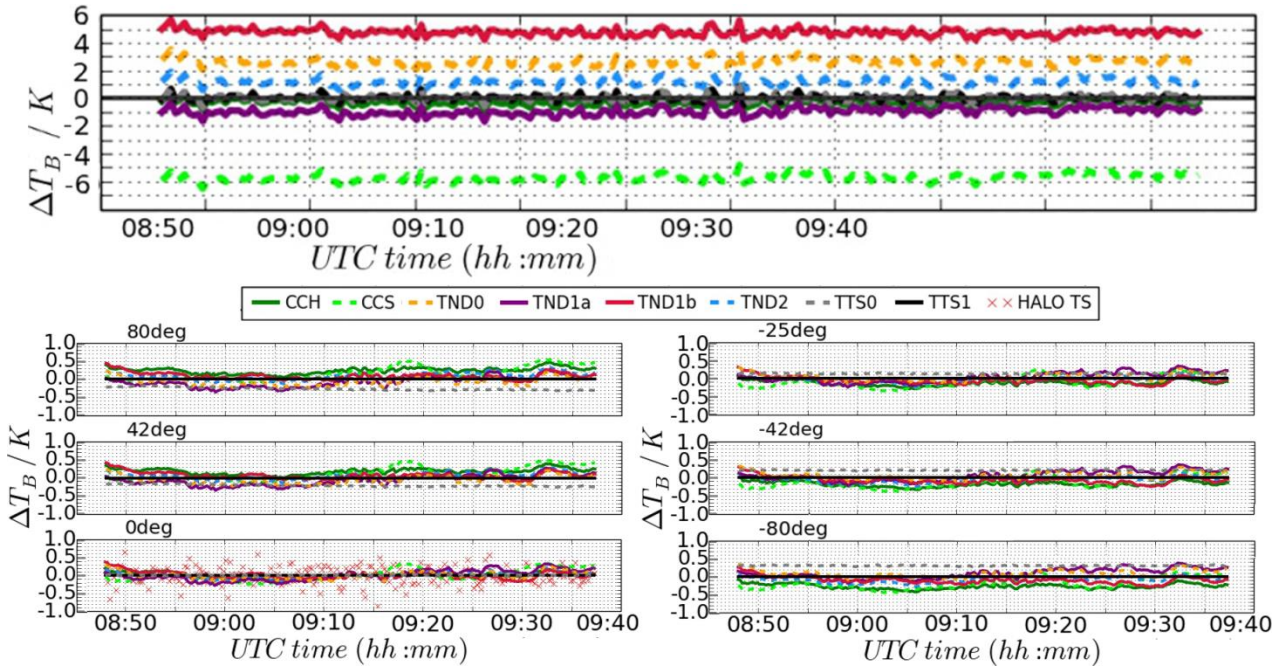


Figure 10: Top panel: Difference between HALO static temperature and TBs derived with the eight calibration methods defined in Table 6 derived from the horizontal measurements at 56.363 GHz during one segment of an ML CIRRUS mission flight. Lower part: Same as top, but plotted relative to method "TTS1", and with applied offset-correction at six different elevation angles measured at 56.363 GHz. Red crosses in lower left panel: difference between method 'TTS1' and HALO TS. Refer to Table 6 for the denomination of the calibration methods.

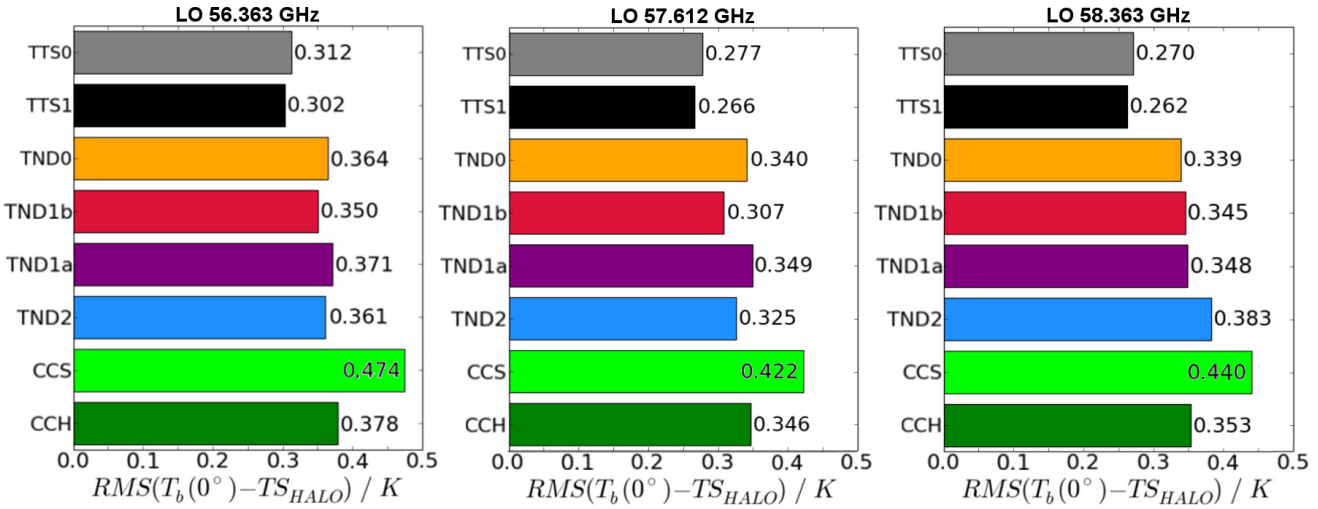


Figure 11: RMS difference between HALO TS and TBs, derived from MTP measurement signal at limb-viewing angle 0◦ at the three standard frequency channels during all ML CIRRUS flight segments with no altitude changes, curves, or ND failures, and longer than 10 min. Refer to Table 6 for the denominations of the calibration methods.

5

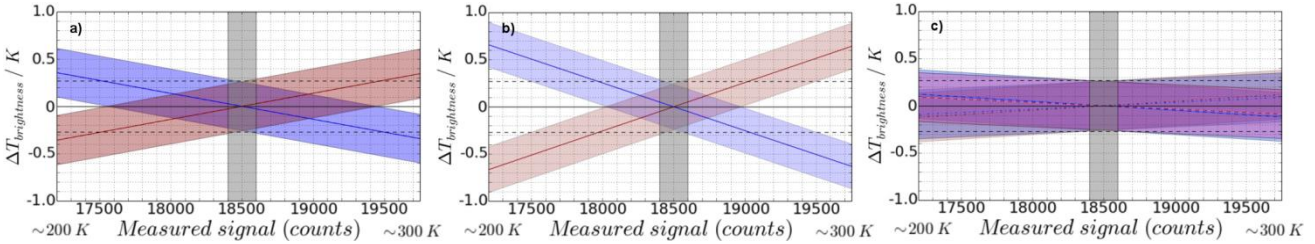


Figure 12: Error estimation of calibration methods with applied HALO TS (assumed to be at 250 K) offset correction for calibration methods a) noise diode + hot target, b) hot target + HALO TS, and c) using cold-chamber calibration parameters. The measurement error for an individual measurement is indicated by the upper-most and lower-most edges of the blue- or red-shaded region (caused by measurement noise). Vertical, grey shaded region: Expected range of measurement signals, if 0° measurement signal is at 18500 cnts (≈ 250 K). Black, dashed horizontal lines: Expected error induced by measurement noise of 6 cnts.

Component	Description	Standard settings
Fairing	Aluminium/ fiberglass to protect the hardware from the environment	
Microwave window	High-density polyethylene, allowing viewing of the atmosphere at ±80° range Groves serve as an anti-refraction 'coating'	
Rotatable mirror	Aluminium mirror ~~designed~~with hyperbolic design to have a beam width of 7.5°; 360° rotatable. ~~Steper~~Stepper motor used: Lin Engineering~~.~~, CE-5718L, step-size: 1.8°	+80°, +55°, +42°, +25°, +12° ±0°, -12°, -25°, -42°, -80°
Horn antenna	Conical, corrugated feed horn with an orthomode transducer (OMT) attached to the base	
Radiometer parts	Crossguide coupler for injection of a noise diode calibration signal Isolator to prevent local oscillator (LO) signal leakage. Double-side-band biased mixer Amplification and an intermediate frequency (IF) filter to select the pass band	Nominal filter band-with: ~ 200 MHz
Frequency synthesiser	Wide band YIG-tuned synthesizer with 1 Hz resolution can be tuned for an output frequency from 12.0 to 16.0 GHz The synthesiser output is doubled twice for a possible LO frequency range of 48 GHz – 64 GHz	56.363 GHz, 57.612 GHz, 58.363 GHz
Reference target	1-inch thick carbon-ferrite mounted on an aluminium plate Styrofoam and ROHACELL foam insulation (1/4 inch) Two conventional power resistors for temperature control; integrated in aluminium plate	Temperature control set to approx~~.~~, 45°C at the back of the target
Data	DC voltage proportional to the brightness temperature in front of the antenna, converted to digital counts	Recorded using LabView software on PC/104

		Integration time (signal recording): 200 ms
Housekeeping	Platinum Resistance Temperature Devices (RTDs) at various locations on MTP hardware	
Controller PC	Commercial ultralow-power single board computer in a PC/104 format with a passive heatsink Runs independent from cabin control	Connected to HALO network to enable user control if necessary or wanted

Table 1: MTP instrument: Components, and settings.

	56.363 GHz	57.612 GHz	58.363 GHz
Gauss fit: lab measurements	Mean: 2.92825e-13	Mean: -1.5046e-12	Mean: -2.64535e-12
(N = 4398)	Std. Dev.: 6.04393 cnts	Std. Dev.: 6.0963 cnts	Std. Dev.: 6.2264 cnts
Gauss fit: flight data	Mean: 2.664e-12	Mean: 9.1452e-14	Mean: 3.7587e-13
(N = 3978)	Std. Dev.: 5.1974 cnts	Std. Dev.: 5.1546 cnts	Std. Dev.: 5.72666 cnts
Auto-correlation (spectral fit)	$\alpha = 0.71$	$\alpha = 0.70$	$\alpha = 0.71$

Table 2: HALO-MTP instrument noise characteristics at each of the three standard frequency channels.

	Scanning unit temperature					Hot target counts c_{hot}				
LO	Ref. T_{sc}	Ref. s_{cal}	Lin. Fit slope	Ref. T_R	Lin. Fit slope	Ref c_{hot}	Ref s_{cal}	Lin fit slope	Ref T_R	Lin fit slope
GHz	°C	K cnts^{-1}	10^{-5} cnts^{-1}	K cnts^{-1}	10^{-5} cnts^{-1}	cnts	K cnts^{-1}	10^{-6}K cnts^{-2}	K cnts^{-1}	10^{-6}K cnts^{-2}
56.363	7.518	0.043154	1.0937	524.492	0.3132	19486	0.043154	2.0141	524.492	0.0647
57.612	7.527	0.040446	0.9989	464.104	0.2716	19292	0.040446	1.8964	464.104	0.0559
58.363	7.474	0.040031	1.7775	492.777	0.4599	20213	0.040031	3.4361	492.777	0.0922

Table 3: Linear fit values linking calibration slope values and receiver noise temperature T_R (calibration *y*-intercept) to MTP scanning unit temperature and hot target counts.

LO [GHz]	ref. $\hat{c}_{ND}$	Ref. T_{ND} [K]	Lin. Fit slope [K $cnts^{-1}$]
56.363	2799	120.90706	0.033089
57.612	3049	123.43799	0.046590
58.363	2932	117.53960	0.052118

Table 4: Linear fit values linking noise diode offset temperature to MTP noise diode offset counts.

LO [GHz]	Ref. T_{sc} [°C]	Ref. T_{hot} [°C]	Lin. Fit slope [°C °C^{-1}]
56.363	7.518	43.271843	0.89126
57.612	7.527	43.036542	0.103719
58.363	7.474	43.211868	0.088969

Table 5: Linear fit values used to correct the MTP hot target brightness temperature.

	Laboratory parameters		MTP hot target + noise diode				MTP hot target + TS	
	'CCS'	'CCH'	'TND0'	'TND1a'	TND1b'	TND2	TTS0	TTS1
Lab s_{cal}	T_{SC}	c_{hot}	-	-	-	-	-	-
Lab T_R	T_{SC}	c_{hot}	-	-	-	-	-	-
T_{ND}	-	-	(u)	(c)	(u)	(c)	-	-
T_{hot}	-	-	(u)	(u)	(c)	(c)	(u)	(c)
TS	-	-	-	-	-	-	(u)	(u)

Table 6: Calibration methods tested with MTP data. T_{SC} indicates linking of the parameters to the scanning unit temperature, c_{hot} indicates linking to hot target measurement signal. Usage of uncorrected data is denoted with a '(u)', applied corrections with a '(c)'.

Error source	Name	Estimation method	Uncertainty	Ref. value
Hot target brightness temperature	ΔT_{hot}	RMS to linear fit in cold-chamber measurements	0.23 K	315 K
HALO static temperature (TS)	ΔTS	RMS to 13s running average	0.13 K	250 K
ND offset temperature	ΔT_{ND}	RMS to linear fit in cold-chamber measurements	0.25 K	120.63 K
Cold-chamber slope	Δs_{cal}^{CCh}	RMS to linear fit in cold-chamber measurements	8.224×10^{-5} K cnt^{-1}	0.04121 K cnt^{-1}
Cold-chamber Y-intercept	ΔT_R^{CCh}	RMS to linear fit in cold-chamber measurements	1.205 K	493.79 K
Measurement noise	Δc	Deviation from linear fit in stable flight segments	6 cnts	18500 cnts

Table 7: Individual uncertainties of values used in brightness temperature calculation.

Appendix: Simple approaches to increase the range of sensitivity

The altitude range of sensitivity and the vertical resolution of the retrieved temperature profile from MTP data depend on the set of frequency channels and the set of elevation angles used when recording MTP data, respectively. For an in-depth test of the optimal settings a full retrieval feasibility study would be mandatory, which is beyond the scope of this study. In the following the results of idealised radiative transfer (RT) simulations (cloud-free, not considering any special cases) are summarized to demonstrate the impact of these settings.

For this assessment the transmission and weighting functions (e.g. Ulaby et al., 1981) are of central importance. Transmission, $\mathcal{T}(\nu) = \exp(-\tau(\nu))$, characterizes the ratio of outgoing to incoming radiation traversing an atmospheric layer with path coordinate s. It is expressed through the optical depth $\tau(\nu)$ defined as the integral of the absorption coefficient (α), which depends on the frequency (ν), the (path-dependent) atmospheric pressure ($p(s)$) and temperature ($T(s)$) of the layer within the plane-parallel atmosphere:

$$\tau(\nu) = \int_0^s \alpha\big(\nu, p(s'), T(s')\big) ds' \quad (Eq. A.1)$$

For brevity, in the following, the path-dependency of α is expressed as $\alpha(\nu, s)$. To investigate the range of sensitivity it is useful to calculate the signal contribution from each respective layer of the atmosphere, determined by the weighting function (WF) defined as:

$$W(\nu, s) = \frac{\partial \mathcal{T}(\nu, s)}{\partial s} = \alpha(\nu, s) \cdot \exp\big(-\tau(s)\big) = \alpha(\nu, s) \cdot \exp\left(-\int_0^s \alpha(s') ds'\right) \quad (Eq. A.2)$$

For RT calculations the Python scripts for Computational Atmospheric Spectroscopy (Py4CAtS[1]; Schreier et al., 2019), a re-implementation of the Generic Atmospheric Radiative transfer Line-by-line IR Code GARLIC (Schreier et al. 2014), written in Fortran, are used. The WFs were computed from absorption coefficients using spectroscopic line parameters from the high-resolution transmission molecular absorption database (HITRAN; Rothman et al., 1998), assuming a midlatitude summer atmosphere (Anderson et al., 1986).

The WFs for the three standard frequency channels used by the HALO-MTP under the nine non-horizontal viewing angles of the standard measurement strategy and assuming an aircraft altitude of 11 km are shown in Figure A1 (left panel). The standard MTP WFs do not show any peaks away from the flight level, indicating that most information is gathered at the aircraft altitude. Nonetheless, from the difference between measurements under varying elevation angles and using different frequency channels, information on the vertical temperature profile can still be

[1] available at ~~http~~https://atmos.eoc.dlr.de/tools/Py4CAtS/

gathered. However, the weights at ±2 km distance to the aircraft are less than a tenth of those close to flight level, indicating that not much information is ~~gathered at~~collected from this distance or further away. At lower altitudes, with higher pressure leading to less transmission beneath the aircraft, this distance is even smaller.

A.1 Choice of LO frequencies

Logically, the best idea to widen the range of sensitivity would be to use different frequency channels that are located at weaker absorption lines than the standard frequency channels, on the wing of the 60 GHz oxygen absorption complex or even between two lines, as was done with the older MTP instruments. However, the choice of an LO at the centre frequency of an absorption line has several advantages: (i) the symmetrical line shape makes the retrieval more exact, (ii) synthesiser errors (small derivations of the LO from the intended frequency) cannot lead to large errors (opposite to a placement in which a strong line is placed just outside the filter range), and (iii) pressure broadening has not as strong an effect as with a placement between two lines. Concerning the threshold of possible frequencies, water vapour absorption becomes important in RT, whenever frequencies close to 50 GHz are used.

To test the influence of opacity of the atmosphere, radiative transfer calculations were made in which the temperatures of all atmospheric layers between ground and 110 km altitude were set to 250 K. The simulation is made using the Temperature InveRsion Algorithm for MIcrowave SoUnding (TIRAMISU~~(~~; Xu et al. 2016), a retrieval algorithm developed to process MTP brightness temperatures, which uses the radiative transfer model GARLIC. Simulations are made for the whole spectrum of frequencies between 50 GHz and 60 GHz with 0.01 GHz resolution. This range includes the three standard frequency channels already in use, but also eight weaker absorption lines (Liebe et al. 1992). Furthermore, the simulations were carried out assuming six different flight altitudes between 2 km and 15 km, which is the ceiling altitude of the HALO aircraft. In this setup, the expected brightness temperature for the horizontal and up-looking viewing directions is 250 K, unless the optical thickness of the atmosphere is small enough that the cold cosmic background is influencing the measurement, leading to a smaller brightness temperature. The more transparent the atmosphere is at any frequency, the colder is the simulated brightness temperature, and the atmosphere close to the aircraft only contributes to a smaller part of the measured signal.

The resulting brightness temperatures are shown in Figure A2. The left panel shows those at limb-viewing angle 0° (horizontal viewing direction) and the right panel those at +80° (near-zenith). The solid, black line in the left panel of Figure A2 shows that for any frequency channel below 54 GHz the atmosphere becomes partly transparent even at the horizontal viewing angle. Hence, those measurements cannot be calibrated (or offset-corrected) using HALO TS, indicating~~,~~ that only frequency channels above 54 GHz should be considered. The results for the near-zenith

measurements (right panel of Fig. A2) indicate that the atmosphere is partly transparent for all possible frequency channels at nearly all flight altitudes. Whenever this transparency is too strong, the signal measured at weak absorption lines while looking downwards could be dominated by the surface temperature, which might not be well-known. As a result, for adding LOs to the MTP measurement strategy, only three possible frequency channels are considered: Those corresponding to the oxygen absorption lines at 54.671 GHz, 55.221 GHz and at 55.784 GHz. The weighting functions of those three possible LOs under the standard set of elevation angles are shown in Fig. A1 (right panel). Obviously, the new frequency channels at weaker oxygen absorption lines are sensitive to a much wider range of altitude layers, especially below the aircraft. However, above the aircraft the weighting functions look similar to those of the standard frequency channels. This is due to the partial transparency of the atmosphere at these frequencies, indicated by low TBs in Fig. A2, combined with the fact, that the viewing direction points through a medium that becomes optically thinner with increasing distance to the sensor.

A.2 Choice of elevation angles

When discussing the choice of the set of elevation angles to be used in the MTP measurements, the signal path through the atmosphere has to be considered. By hardware-design limitations, the range of MTP viewing angles is limited to ±80°. To consider a new, feasible set of elevation angles, it makes sense to compare the path lengths of all possible elevation angles α with the shortest possible path length through a vertical layer of the atmosphere at maximum elevation (±80° relative to the horizon):

$$l_{rel80^\circ} = \frac{\cos(10^\circ)}{\cos(90^\circ - \alpha)} \quad (Eq.A.3)$$

The relative path lengths to the ±80◦ angle are summarised in Table A1. Especially the three largest elevation angles used in the standard MTP measurement strategy (underlined values in Table A1) do not differ much in their path lengths. This can result in the WFs of different measurements being very similar (overlaying lines, e.g. below aircraft altitude in Fig. A3); those measurements are (partly) redundant.

To derive a new set of elevation angles for MTP measurements with as much independent information as possible, a rule of thumb is used, that with each new angle the length of the signal path at 80° should be added, meaning that l_{rel80° is close to an integer. Corresponding rows are highlighted in grey in Table A1, including one angle with $l_{rel80^\circ} \approx 1.5$. However, due to the fact that the antenna beam of the MTP instrument has a field of view of 7° - 7.5°, the measurements at 11° and 14° would overlap, and probably also not contain much different information from the measurement at 19°.

A.3 Determining a new measurement strategy

Since the MTP is mounted on a moving platform with approximate speed of 200 m/s, it is also necessary to consider the time it takes to record one complete measurement cycle. In favour of better horizontal resolution, the most appropriate set of elevation angles is ±14°, ±30°, ±41°, and ±80°, ~~taking into account~~considering the field-of-view of the antenna. Thus, including the horizontal measurement, only nine elevation angles would be used, instead of the 10 standard angles, in which the down-looking set of angles is smaller than the up-looking set; leaving out the -55° limb-angle. Since the up-looking WFs of all possible frequency channels are very similar, the opposite would be more feasible: using more down-looking angles to enhance the resolution of measurements below the aircraft, but reduce the number of up-looking angles, e.g. by leaving out the +41° measurement.

Based on all previous considerations, four new measurement strategies are proposed and summarised in Table A2. The new strategies are compromises between vertical resolution and range of sensitivity, keeping the total number of measurements per cycle close to the original, so the total time of recording a complete measurement cycle does not change significantly, keeping the horizontal resolution of measurements. All proposed strategies use eight viewing angles and four frequency channels to enhance the vertical resolution and altitude range at the same time. The weighting functions of the measured signals for each of those new strategies are shown in Fig. A3. Depicted are the cumulative weights to indicate the percentage of the measurement signal that is acquired with ~~increasing~~decreasing distance to the aircraft. This depiction helps to understand how much a certain layer of the atmosphere contributes to the total signal at a single viewing angle and frequency. If two lines in the figure over-lap, the corresponding measurements (i.e. measurements at two certain frequency- and viewing angle combinations) are redundant. To compare the relative contributions of different frequencies to the total incoming signal, please refer to Fig. A1.

Strategy ‘8E4LOa’ shows the result of simply adding a frequency channel to the standard set (Fig. A2a). In the other three proposed strategies, only two frequency channels of the original set are kept, and two frequency channels at weaker absorption lines are added. Since it is desirable to have the least redundancy in the measurement, overlaying weighting functions, as seen, for example, in Fig. A2 d) are to be avoided. Also, if the aircraft is flying at lower altitudes, the frequency channel corresponding to the weakest absorption line might be influenced by the surface temperature. Hence, depending on the planned flight pattern, strategies ‘8E4LOb1’ (Fig. A2c)) or ‘8E4LOc’ (Fig. A2b)) should be favoured. The simple approach taken here, indicates that both strategies increase the MTP measurement sensitivity to a wider altitude range, especially below flight level, while allowing for a well-resolved temperature retrieval and keeping the horizontal resolution of MTP data.

Obviously, the effects of changing the measurement strategy depend on the atmospheric conditions, mainly the true temperature profile around flight altitude. Hence, further investigations using forward radiative transfer calculations ~~that are taking into account~~considering all MTP instrument characteristics shown in the main text, would be needed to determine the influence of the proposed changes on retrieval input error, as well as vertical measurement resolution. As the outcome clearly depends on the chosen retrieval algorithm, this is to be done in a separate study, related to the retrieval algorithm used.

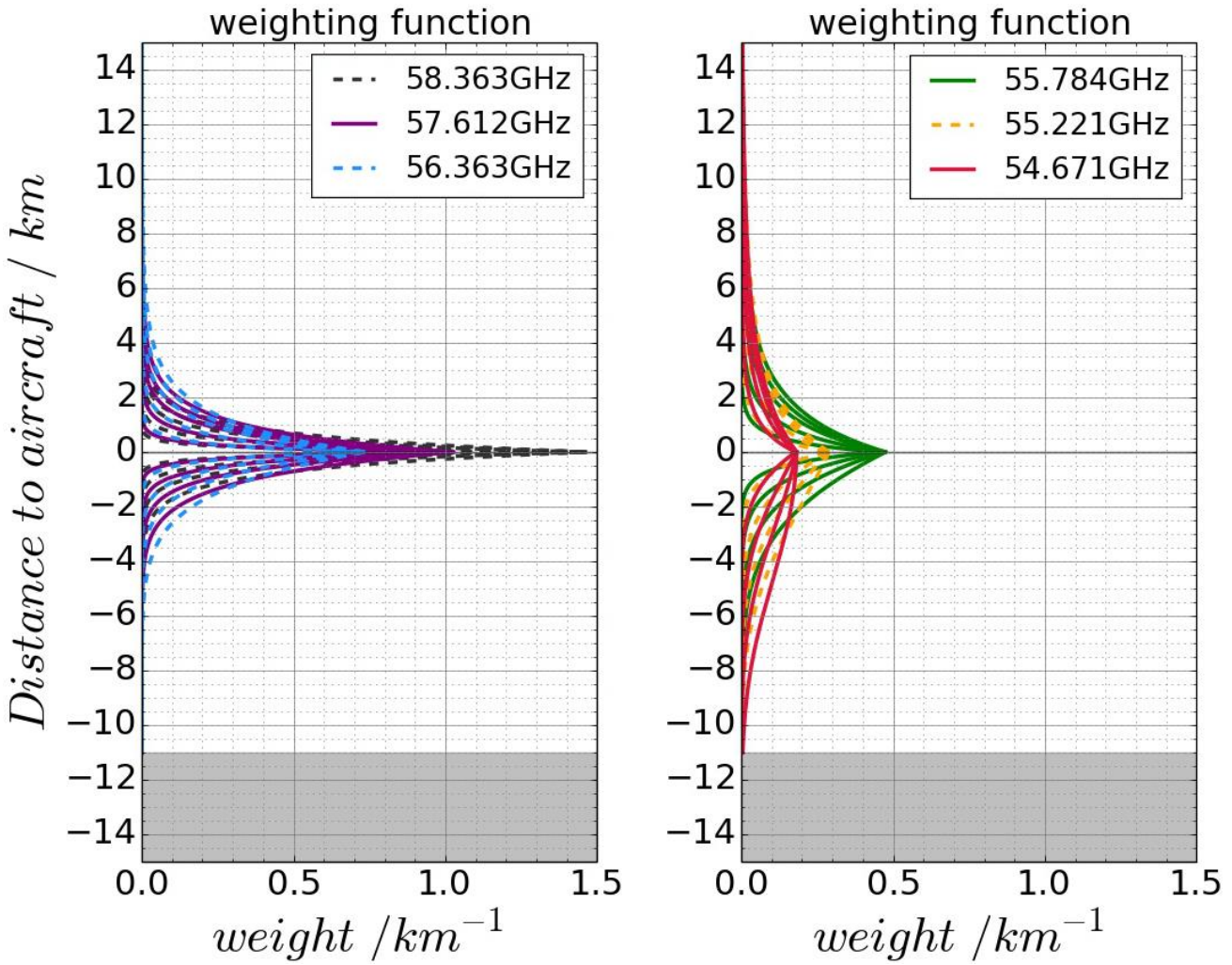


Figure A1: WFs (averaged over all contributing frequencies within filter transmission range) of the MTP frequency channels (each individual line corresponding to a different viewing angle), calculated at aircraft altitude of 11 km. Shown are the three standard frequency channels (left), and three possible frequency channels (right) to be considered in a new measurement strategy of the HALO-MTP. Grey areas at the bottom: altitude range that would be below the surface. Note that the black curves for the 58.363 GHz measurements are almost entirely covered by the other lines.

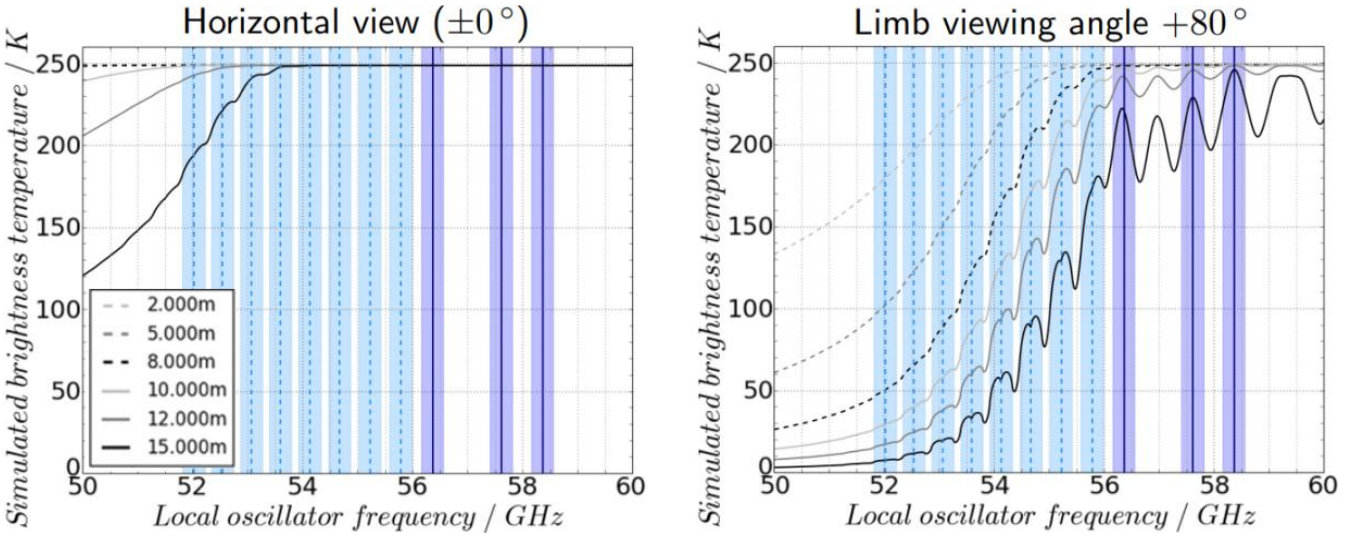


Figure A2: Simulated TBs at frequency channels between 50 GHz and 60 GHz at different flight altitudes (different line styles) and at horizontal viewing angle (left panel), and at near-zenith angle (right panel). Solid, vertical lines: Standard frequency channels of the MTP; dashed vertical lines: strong lines that could be used as new MTP LO. Shading around vertical lines: MTP filter width. Lines corresponding to altitudes of 8 km or lower overlap in the left panel.

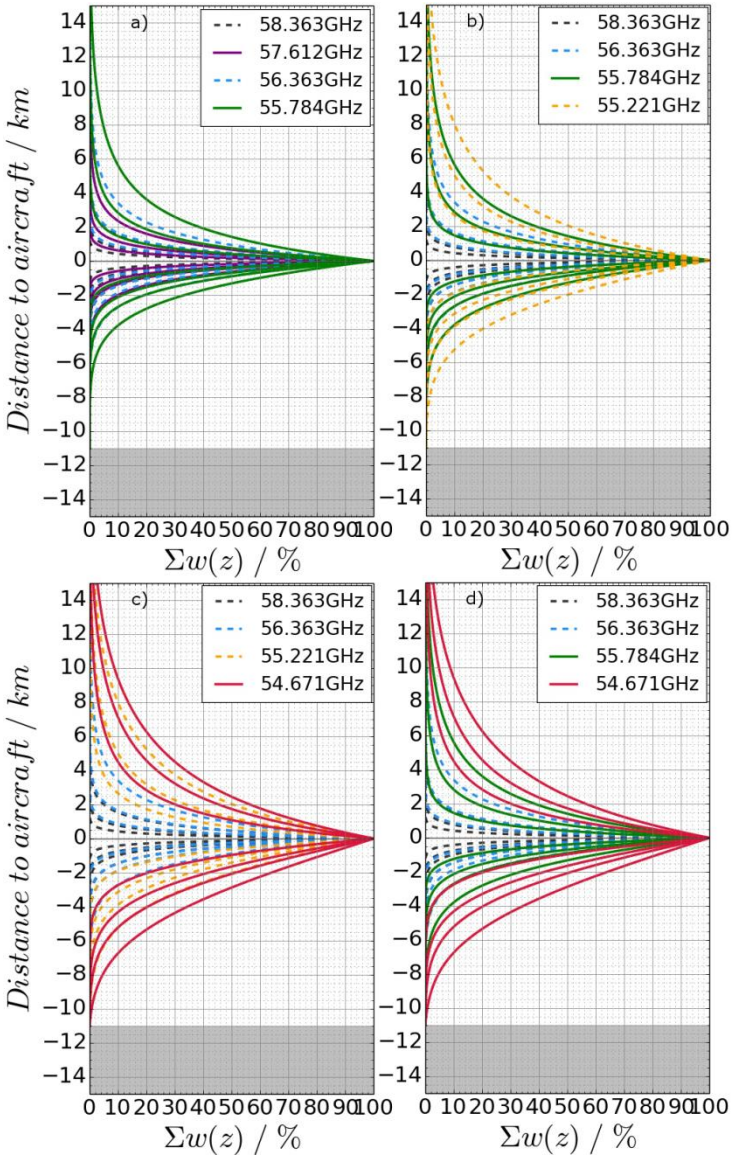


Figure A3: Cumulated weights of the MTP frequency channels (each individual line in a panel corresponding to a different viewing angle and frequency combination), calculated for an aircraft altitude of 11 km. Shown are possible new measurement strategies, as mentioned in Table ~~A1~~A2 a)'8E4LOa' b)'8E4LOc', c)'8E4LOb1', and d)'8E4LOb2'. Grey areas at the bottom: altitude range that would be below the surface.

α	$l_{rel80°}$	α	$l_{rel80°}$	α	$l_{rel80°}$	α	$l_{rel80°}$	α	$l_{rel80°}$
1°	56.428	13°	4.378	19°	3.025	28°	2.098	41°	1.501
[…]	[…]	14°	4.071	20°	2.879	29°	2.031	42°	1.472
5°	11.299	15°	3.805	[…]	[…]	30°	1.97	[…]	[…]
[…]	[…]	16°	3.573	25°	2.33	31°	1.912	55°	1.202
11°	5.161	17°	3.368	26°	2.247	[…]	[…]	[…]	[…]
12°	4.737	18°	3.187	27°	2.169	40°	1.532	80°	1.0

Table A1: Signal path lengths relative to ±80°. Underlined: Elevation angles used in the standard measurement strategy. Grey cells: Possible candidates for a new strategy. Elevations in between (marked by '[…]') do not correlate with integer values in relative path lengths.

Name	Elevation angles	LOs [GHz]	t_{cyc}
standard	+80°, +55°, +42°, +25°, +12°, ±0°, -12°, -25°, -42°, -80°	56.363, 57.612, 58.363	~13 s
'8E4LOa'	+80°, +30°, +16°, ±0°, -16°, -30°, -41°, -80°	55.784, 56.363, 57.612, 58.363	~14 s
'8E4LOb1'	+80°, +30°, +16°, ±0°, -16°, -30°, -41°, -80°	54.671, 55.221, 56.363, 58.363	~14 s
'8E4LOb2'	+80°, +30°, +16°, ±0°, -16°, -30°, -41°, -80°	54.671, 55.784, 56.363, 58.363	~14 s
'8E4LOc'	+80°, +30°, +16°, ±0°, -16°, -30°, -41°, -80°	55.221,55.784, 56.363, 58.363	~14 s

Table A2: Proposed measurement strategies for future missions of the HALO-MTP.