I: Author responses to Reviewer comments

Reply to the referee #1 comments

We would like to thank the referee for thoroughly reviewing our manuscript and for the helpful advice provided in the comments, below, which we believe helped 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 referees 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

General Comments:

1. The implications of the instrument characterisation for the subsequent interpretation of TB and temperature retrieval are not thoroughly assessed. Section 5 should have a clear outcome on the questions: (i) spectral characteristics: Which are the representative frequencies of the three channels? Which frequencies shall be assumed for the retrieval algorithm? Does the RT have to consider the full bandpass characteristics? (ii) Which is the effect of the antenna bandwidth? Is a pencil beam approach justified? (III) What noise characteristics have to be assumed in the retrieval, e.g. in the measurement covariance matrix?

This is a very thorough list of characteristics to be assessed for a retrieval algorithm. Currently, there exists more than one approach to retrieve the absolute temperature profiles from MTP brightness temperatures, which is one of the reasons why the authors decided to not consider retrieval algorithms in the study.

The intention of Section 5 was to point out that a different approach to decide which measurement strategy (i.e. number of viewing angles and LO frequencies) could be used. A thorough study of all implications for a retrieval algorithm would be beyond the scope of this study. For that reason and on advice of the reviewer, Section 5 has now been moved to an appendix. The usefulness of an assessment of the questions asked above, as well as to investigate the impact of changes in the measurement strategy are understood, and should be attempted in an upcoming study.

2. Accurate calibration is the most important task in microwave radiometry. As all components are strongly temperature dependent besides temperature stabilisation a periodic calibration is needed. The calibration might only update the gain of the system (relative calibration) or make an absolute calibration in which all parameters of the raw measurement (count) to TB model are derived. In the simple linear case (as it is used in this manuscript) these are gain and receiver noise temperature Tr which can be derived by pointing the antenna successively to two reference targets. The authors seem to be not aware of this classical microwave formalism which is also apparent as they hardly cite any literature microwave radiometry (list in the back) and some flaws in the radiometer formula application. The major questions which would need to be ad dressed are: How good are the reference targets (blackbodies)? How frequently does a calibration need to be made? Why have the measurements in the cold chamber with view on a stable target not been used for such an analysis? The next step would then be the in flight calibration. Assuming that the laboratory calibration (strategy 1) would work is a

bit naïve. However, there are good approaches later on using the horizontally pointing measurements but a motivation and explanation why this procedure was chosen needs to come first.

We think it is not necessary to explain the standard method of radiometer calibration, here. The presented assessment of calibration methods is based on using this calibration method, since it was used throughout the measurement series inside the cold chamber, and the derived calibration parameters are used to characterise various parts of the instrument, and calibration methods used in flight. As pointed out within the manuscript, the instrument only has one single built-in calibration target, which can be used during flight, and the study provides a guideline, how this can be used to derive the brightness temperatures from the measurements taken during flight.

As for the specific questions asked by the reviewer, we have added specific information about the microwave absorber used in the laboratory calibration, as well as some information from the instrument documentation.

Concerning an analysis of the necessary frequency of calibration, please see our comment on the specific comment about an assessment using the Allan Variance, below. In general, the deployment of the MTP instrument, which is mounted completely outside of the aircraft, requires permanent monitoring of the instrument state, which is already implemented in the way the measurements are taken: Calibration measurements are made after each measurement cycle (i.e. after taking measurements at each elevation angle), approx. every 13-14 seconds. This is the highest possible frequency with which calibration measurements can be taken during flight.

Finally, while the proposed strategy no. 1 for in-flight calibration might sound naïve, we have included this approach in our discussion for reasons of completeness. The results show that using this method in the cold-chamber provides the possibility to link in-flight instrument conditions to measurements within the cold-chamber, and derive calibration parameters, that lead to comparable results as other calibration methods.

In the revised manuscript we therefore have included a brief discussion of the pros and cons of using each of the calibration strategies when introducing them.

3. *The information on the MTP measurement principle is not clearly provided in the beginning of the manuscript making it difficult for the reader to follow. Bits and pieces come together at different instances, e.g. scanning is explained on page 14 and especially the discussion on the use of different oxygen lines is confusing. For better understanding the authors should include a thorough description of the MTP measurement principle in the beginning and add an absorption spectrum (preferably even for different pressure levels as in Fig. 16) to illustrate the frequency channels (and their potential tuning range). This also serves to introduce the double sideband principle. Further, it could be explained why the LO is typically set at center frequency for mitigating problems due to frequency drifts, and how non-resonant emission (water vapor continuum, hydrometeors) affects the measurement. This would also demonstrate that the LO frequency is not the frequency for which the measured TB is representative (passband averaged – see Fig. 16).*

The revised manuscript now contains more information on how the measurement set-up works, and links to the already existing literature, in which the standard settings for this instrument are already introduced and discussed (i.e. Mahoney and Denning, 2009, and Lim et al., 2013). We have updated Table 1 to include all relevant information.

The measurement principle is explained in more detail in the revised version.

4. *Section 5 address future measurement strategies in terms of frequency selection and elevation scanning. This is an important study but is not done as thoroughly as it is needed especially in light of vertical resolution of the retrieved temperature profiles for different types of atmosphere. It also does not take into account the findings of*

their laboratory measurements in respect to the spectral and spatial sensitivities. As the paper is already very lengthy it should be taken out.

We agree that more could be done to thoroughly assess all input parameters needed in a retrieval. We are aware that those input parameters may depend on the type of retrieval used (optimal estimation, neuronal network, Tikhonov retrieval,…), and the Section was intended to indicate potential for possible improvements in the general measurement strategy. While we do think that this consideration is worthy to be noted by users of the MTP data, and other groups using MTP instruments, we agree, that it might better fit as an appendix.

5. At several instances it seems that the authors have gravity wave detection as application in their mind – this is ok but needs to be clearly stated (only abstract). Many readers might not know which requirements in TB are needed for this purpose. Other users might be more interested in vertical resolution for stability assessment.

Indeed, the study was first undertaken with the goal of assessing temperature fluctuations for gravity wave studies. As the short overview of already published studies using MTP data shows, there are other interests in using this kind of data, as well. We have changed the text in the abstract and introduction accordingly.

6. The readability of the paper needs to be improved - sometimes it is more a technical report than a paper. No clear goals are provided, the structure is not always clear, the text is written rather lengthy and many basic informations only appear rather late in a middle of a section where you would not expect it. Short paragraphs sometimes even only one sentence long occur and the text frequently repeats (unnecessarily) the figure captions, e.g. "Plotted is also a..". The paper could be shortened by reducing number of figures or using an appendix. I would recommend to concentrate only on the past measurements. The optimized scanning strategy In case but the future – which I think would be an own study if done carefully could go in an appendix.

As noted above, we agree, that section 5 fits better as an appendix to the study, and have changed the manuscript accordingly. The wording of the manuscript has been checked, and where necessary, revised.

Figure 8 in the revised version now contains Figures 8 and 10 of the submitted version, and Figure 10 of the revised manuscript now contains both Figures 11 and 12 of the submitted version.

Specific Comments: Why are brightness temperatures referred to as BT in the text (and Fig. 11) and TB in the equations. Historically the satellite community uses BT and the ground-based community TB. I don't think it matters which one is chosen but it should be consistent.

All instances of "BT" in the text have been changed to "TB" for consistency.

P1l8: "records radiances", no it records counts which are calibrated to brightness temperatures - it is ok to say TB here

Indeed, the recorded signal is in counts. Since the physical quantity that is measured is radiance, the sentence is changed to: "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."

P1l9: "state of the atmosphere can be derived" this indicates much more information than the temperature profile which was stated already – what else?

3

The (dynamical) state of the atmosphere can be derived from the temperature profile; The sentence is extended to: "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)."

P1l22: "weaker oxygen lines" better write 'frequency channel'. The LO frequency of the channel does not necessarily need to be at a line center. Also and it seems to me that it is not clear to authors: the LO frequency is not the representative frequency of the channel – and the "representative frequency" can be extracted from their laboratory measurements. I anyway suggest to modify section 5 such that it can provide the necessary input for the retrieval algorithm

In the case of the MTP, the LO frequency really is placed at the line centre, which is stated in section 2.2. The discussion of the "representative frequency relates to the radiative transfer calculation, which is part of the retrieval algorithm. The transmission function was measured for the MTP, and should be used when setting up the retrieval calculations. In section 5, a simplified approach was chosen, to show that some easy-to-be-made changes to the measurement set-up can already have a large influence on the quality of retrieval input. Since this section will be shifted to the appendix, we will not attempt to make additions, as a thorough investigation of implications for radiative calculations may well fill its own study.

P1l22: "calibration parameters do clearly depend on the state of the instrument". This is the key in microwave radiometry for astronomy, atmospheric, planetary science etc. ever since and for all instruments there is the question how frequently one has to calibrate, e.g. Dicke switching for short-term fluctuations. Unfortunately, even slight vibrations and temperature changes can cause transmission characteristics to change thus calibration parameters. So this sounds a bit naive – I recommend the authors to look more in basic microwave radiometry books, e.g. Janssen, 1994, Vowinkel, 2013, Woodhouse, 2017,

We agree that this problem is not unique to the MTP. Microwave instruments always have this issue, and the literature states many examples and approaches used to stabilize the systems in operation. Since the MTP is mounted completely outside an aircraft, which frequently changes flight altitude and/or speed, and enters different temperature and wind conditions, such stabilization is not readily possible. Moreover, changing the instrument design in an attempt to increase instrument stability would be very costly, and the limited space inside the wing canister does not allow for many options in hardware changes anyway. This increases the problem for the MTP, as compared to, e.g., ground-based instruments. We have changed the sentence to show these circumstances: "The MTP shows quite large changes of the instrument state, imposing considerable changes in calibration parameters over the course of a single measurement flight"

P1l26: Here it should be said that precision is determined for TB which closely relates to the atmospheric temperature when the instrument is pointed horizontally – otherwise it is confusing

We have made the suggested change.

P2l16. What is meant by structures?

Basically it is meant that any signal in the timeline of measurements, that deviates from a smooth background could be caused by either noise created by the instrument itself, or has its cause in some real, physical process in the atmosphere, in which case we would like to be able to detect it with this instrument.

The word "structures" was replaced by "fluctuations".

P3l9-12 and P3l14: There is a very long list of applications of past studies using older versions of the MTP (is that really necessary?) and then it is claimed that instrument characteristics need to be known for correct interpretation. This is true and that's why this study is valid but it somehow implies that the work here also helps with data from old campaigns. This needs to be clarified.

It was not the authors' intention to claim that this study helps interpreting measurements from older MTP instruments. Listing the previous usage of MTP data from the past was intended to show that data from MTP instruments developed by JPL is frequently used and that this study is of interest to a wider community of researchers.

The paragraph has been reworded so that there is no impression given that older studies may not be valid due to the herein presented findings. In the manuscript we only state that thorough characterisations of older MTP instruments have not been published before. We are clearly not implying that those characteristics were always completely unknown!

Introduction: the whole introduction is dedicated to the MTP but there is no reference to other studies on the characterisation of other microwave airborne instruments is made, e.g. Blackwell et al, 2001 describing NAST with frequencies 50-57 GHz, McGrawth and Hewison, 2001, Wang et al, 2007 etc. which might also check different instrumental parameters. The introduction clearly needs to mention the goals of the lab investigations.

The studies mentioned by the reviewer provide a variety of instrument characteristics, many of which were also referred to here. Other characteristics measured, e.g. in McGrawth and Hewison were not an option for the DLR.MTP, as a disassembly of the instrument hardware was not possible without losing aircraft certification. This is now stated in the introduction and in the appropriate sections, in connection with a reference, where fitting.

P4l2: Not all radiometers for temperature profiling measure at the oxygen absorption complex around 60 GHz - also 118 GHz is used. In general, it is surprising that no reference is made to the fact that operational meteorological satellite instruments, e.g. AMSU-A, do temperature sounding since decades. These sounders exploit only the frequency information for profiling while the MTP aims at improving the resolution by angular information. It is necessary to explain the measurement principle here thoroughly, showing a spectrum (ideally for different altitudes) and the considered frequency channels. On a side note: The accuracy of the oxygen spectroscopy is still under debate which is, however, more important for retrievals, Caddedu et al, 2007; Cimini et al, 2018, Maschwitz et al 2013.

It is true, that this study places its focus on the MTP instrument itself. We have changed the wording to acknowledge the fact that temperature measurements are also possible at 118 GHz. The comment on using the angular information to improve the vertical resolution is very valuable, and we changed the text in Section 2.1 accordingly.

The development of the MTP instrument has not been done by the authors of this study. Since the instrument has a long history, there are a number of publications available, explaining the measurement principle (e.g. Denning, et al. 1989), as well as some unique features and considerations related to the wing-canister design (e.g. Lim et al. 2013), including consideration of the used frequencies. We have given those references in the

description of the MTP instrument (Section 2.2) and added a few sentences to briefly introduce the measurement principle.

Concerning the very interesting ongoing debate of the accuracy of oxygen spectroscopy, we agree that this topic relates more to retrievals, which has been explicitly excluded from this study.

P4l2: Why don't you explain the heterodyne principle and talk about a double side band receiver. This is very important to clearly define the frequencies for the radiative transfer used for retrieval development.

The important information that the DLR-MTP uses a double-sideband heterodyne receiver is added in the text, and the measurement principle is briefly explained in section 2 of the revised manuscript.

P4l18: "making the retrieval of temperature profiles possible" Most instruments only use information on frequency dependence. Make clear that the MTP can achieve higher vertical resolution by adding the angular information.

We agree, that it is valuable to pointing out a clear advantage of the MTP instrument in comparison to other microwave systems. We have revised the text, making this point clearer (also according to the comment referring to p4l2).

P4l24: Thermal stabilisation is the most important part in a microwave radiometer the performance of all microwave components strongly depends on temperature. Therefore more details on that are needed.

Details on the temperature control are given in Mahoney and Denning, 2009, to which we refer at the beginning of the Section. According to this publication the "[t]emperature control at the point where the thermistor is mounted is approximately ±0.1C".

There is no possibility to monitor the real temperatures of the components during flight, other than through the housekeeping data, recorded during flight. Here, we do see changes depending on flight levels (surrounding temperatures), and temperature gradients across the instrument are visible. However, since the thermistors are only placed at certain positions, there remains the question, if the temperature recorded in the MTP housekeeping data is representative for the critical components, and how changes are to be interpreted.

We do acknowledge that a new series of cold-chamber laboratory measurements to investigate the overall temperature behaviour of MTP components could be an interesting study in the future, but such measurements cannot be performed in the near future to be included in this study.

P4l229: What about temperature stability, homogeneity, spill over of the target, cf. Mc Grawth and Hewison (2001).

While we were not able to perform all of the characterisations described in the study published by McGrawth and Hewison, most tests were also performed in our laboratory set-up.

The main reason for leaving out some of the tests, e.g. determining the spill-over, is due to the fact, that the DLR-MTP is certified to be flown on a research aircraft. Disassembling the instrument would lead to a costly process of re-certification, which prevented us from taking any parts out of the instrument. Hence, a spill-over measurement for the antenna was not possible. For the same reason, we do not have the means to add any observation

system to the instrument while mounted on the aircraft, to, e.g. monitor the thermal stability during flight, or the characteristic temperature of the heated target.

The temperature stability of the target is part of the investigation presented in Section 4. The housekeeping data only state the thermistor temperature in the back of the target, other measurements are not available. Testing the representative brightness temperature of the hot target in the cold-chamber is the nearest we can cat to such an assessment. There is no possibility to conduct an assessment comparable to that shown in McGrath and Hweison, 2001, since we cannot change or add parts to the hardware of the instrument, without a costly revision of the permit-to-fly.

P5, l14-15: the discussion on the oxygen spectrum and LO needs further explanation and should come before not in the section on wing-canister, same for the information on the frequency range (l25) below.

The Section is structured in a way to introduce the basic principle of Microwave radiometry, moving from the broad principle to more and more specific details.

As mentioned in a previous answer, the discussion of the Oxygen spectrum and choice of LO is already presented in Mahoney and Denning, 2009, and also briefly discussed on Lim et al., 2013. We have added a sentence to the text, which points out that the original choice of standards LO frequencies was made by the inventors of the instrument.

P5l22: how large is the gap, x MHz?
The gap is nominally 20 MHz wide, which is confirmed by the measurement of the filter function (Figure 2, left panel). We have added this information in the text, as well as a reference to Figure 2.

P6l 6 "investigation OF calibration"

We have corrected this typo.

Section 3. The frequency response of the bandpass is investigated but there is no discussion on the stability of the LO frequency – does this have any potential effect on measured TB?

A measurement of LO frequency stability would require disassembling part of the instrument, to attach the measurement equipment (oscillator) to the frequency synthesizer output. We did not attempt such measurements, which required disassembling the instrument, as this would have had serious implications for continued airworthiness.

Some thoughts on the potential influence of measured brightness temperature, influenced by synthesizer errors: Potentially the largest influence is that with a shift of the LO frequency, the gap in the middle of the filter function is no longer located at the center of the strong oxygen absorption line. Hence, on one side of the filter function, much larger absorption very close to the aircraft would be included in the signal, while on the other side, the signal is caused by absorption slightly further away from the aircraft. Since the LOs are placed at very strong absorption lines, the first affect probably outweighs the second effect, so that the measured brightness temperature would be representative of an altitude layer closer to the aircraft, than assumed. The absolute error depends on the aircraft altitude and the temperature gradient present in the atmosphere surrounding the aircraft. Largest errors would certainly be induced at lower pressure, where the line shape is sharper. Also, larger

temperature gradients in the atmosphere would induce a larger error in the measured brightness temperature. The absolute effect of frequency shifts in the LO would have to be modelled, using radiative transfer calculations. However, we feel that the large gain fluctuations seen in the calibration of campaign data can be assumed to be much larger than the induced error by small LO frequency shifts.

Following the comment from Reviewer 2, that the pointing error seems misplaced in the Section on brightness temperature calibration, we have added a new Section to discuss further sources of measurement uncertainty, in which the pointing error is included, as well as some discussion of other error sources, such as synthesizer errors, reflecting the discussion above.

P6l27: The authors mention the periodicity of the signal first. I understand that for gravity wave detection this is important but in terms of radiometer performance the most important question is whether the instrument follows the radiometer formulae (Eq. 4.8), i.e. noise reduces with increasing integration time. For this purpose typically the Allan variance is used. This characterizes the noise and determines how long measurements can be integrated in time and how frequently a calibration needs to be performed.

We indeed measured time series of the Tb in the cold chamber. It was found out that up to averaging times of > 20 s the noise behaviour resembles white noise. The precision of the measurement could thus be improved by longer averaging times, however, the spatial resolution due to the high aircraft speed clearly calls for reducing the measurement cycle to short integration times.

A brief discussion of those points is added in the revised manuscript.

Section 3.1: The name is irritating as it could mean much more. The measurements of the bandpass characteristics and the antenna diagram (section 3.1) are important and interesting but are presented rather briefly without any implications for the subsequent retrieval. Even the exact measured bandwidth and beamwidth are not given. For the analysis or implications RT calculation would play a major role. As for example shown in Crewell et al. (2012, their figure 10) the bandpass characteristics can cause the effectively measured TB being representative for a frequency deviating significantly from the specified channel frequency. In fact in the double side band approach this anyway takes place and needs to be handled in the RT underlying the retrieval process. Similarly, the antenna pattern smears out atmospheric features especially at low deviations from the horizontal in a vertically stratified atmosphere (Meunier et al., 2013). To appreciate this laboratory measurements and their impact on the measured TB further analysis is required which would fit well into section 5.

In Crewell, 2010, the authors state that ".Because of the strong nonlinear changes in brightness temperature with frequency when atmospheric spectral features are measured the detector's exact band-pass characteristics have to be taken into account". Our study presents exactly this transmission function, needed for the correct set-up of a retrieval algorithm. Similarly, the antenna pattern is presented in this study, and should, ideally, be used in any retrieval algorithm used to derive the atmospheric temperature profiles.

We did explicitly state that we do not discuss retrieval methods and associated uncertainties; Hence, we feel that while those are important points to consider in a retrieval set-up, a thorough discussion of these effects would go beyond the scope of this study.

P7l12: "A certain 'waviness' is visible next to this" ripples are typical in any microwave component due to EM wave theory propagation – reducing the amplitude is key.

Thank you for your comment on this observation, clarifying the source of the observed signal. Since the structure is now known through the measurement of the transmission function, it can be considered in RT calculations. Attempting to reduce the amplitude would necessitate the disassembling of the instrument, and possibly replacing parts, which, as mentioned before, has serious implications for aircraft certification.

P7l23: how stable is the noise diode, how much does it depend on temperature (stabilization)?

There is an entire section (4.2) dedicated to this question. We have added a reference to this section. When investigating measurement flight data, it is mentioned, that on top of general noise diode signal dependence on temperature, we did experience technical problems due to a cold soldering joint in our measurement campaign, so it is not possible to make statements about the stability in real flight conditions.

P8l14: "takes some time to stabilize".. needs to be more quantitative – later it is mentioned but not here

Actually, quantification is hard due to the fact that this stabilization depends a lot on the environment, and the way the instrument is operated. We did observe quite different times the instrument took to stabilise – between different operating environments. We have added a sentence in the revised text to acknowledge this fact.

Section 3.2: Information on the accuracy of the target temperatures is missing. P9l14 mentions the "hot" target – should be explained before

The reference "hot target" is introduced at the beginning of Section 3.2, in the paragraph that describes the laboratory settings, and which targets are used. (p.8, l21 in the discussion paper).

The heating of the target is done in the same way as the heating of other parts of the instrument (see Mahoney and Denning, 2009). The thermistors are heating the components to within an accuracy of +/- 0.1C (private communication with Mahoney and Denning)

The temperature actually seen in the calibration process is investigated in Section 4.2. Figure 10 shows a constant reading of the "target temperature" from the MTP housekeeping data throughout the entire measurement series in the cold-chamber. There is no other measurement of the target temperature available, and external monitoring, e.g. through an infrared camera, cannot be realised in flight, due to the aircraft certification process. (See also our answer to your comment referring to p4/l29).

P8l30: I find the term "at all LOs" confusing – also at other instances. Why not write for all frequency channels?

Given that the use of "LO" might be misleading to some readers, we gladly follow this suggestion, to rather use the wording "all frequency channels".

P9l7: Why do the authors not use the classical microwave notation using the gain (cf, Janzen, Mc Grawth and Hewison, ? The difference between receiver and system noise temperature needs to be made clear.

We chose this approach based on the investigation of "how to best calculate the brightness temperature from a known recorded signal (counts)", i.e. looking for a way to calculate the brightness temperature as a function of the recorded counts, contrary to the traditional approach taken in microwave radiometry. However, this approach is still similar in the way that the defined slope of the line ("S_cal") in this study, is the inverse of the traditional

definition of the Gain, while the receiver noise temperature is still defined in the same way, as in the classical formulation of microwave radiometry.

Since the classical notation is much better known, the authors do acknowledge this fact, by adding a note in the revised manuscript.

P9l17: Radiometers are never completely stable which is why periodic calibrations have to be made. In between this calibrations the TB could be corrected assuming a linear trend as shown in Fig. 6. The following paragraph describes this for the airborne measurements bit it is unclear for me that for these linear fits segments of 5 min without calibration are used?

The 5 minutes mentioned in the following paragraph do not refer to the time between two calibration measurements, but to the length of the flight segment from which data is being used. Calibration measurements during flight are part of every single measurement cycle (i.e. one calibration measurement every 13-14 seconds!). We have added this information in the revised manuscript.

P10l1 and following: The spectral analysis is interesting and similar to the Allan variance but is unclear to me why it is applied to atmospheric measurements and not to the cold chamber measurements where the real instrument performance could be tested. The concatenation eliminates real temporal signals. Does the analysis differ between in flight and laboratory measurements .

We used the flight segments, because some of those are much longer than the cold-chamber measurement segments. When comparing the analysis from the flight segments to drift measurements in the laboratory, the results do look similar, as long as the drift period is included. If it is not included, the result from the drift measurements indicate a smaller correlation coefficient alpha, showing that without drift, the instrument noise behaves more like white noise.
Using the campaign measurements has the advantage, that the parameters used to test significance in the data analysis represent much more conservative limitations, so that the confidence in the results is higher.

P10l20-27: "line parameters" is irritating as it could be interpreted in spectral lines: it is about the updating your calibration model, basically, gain and receiver noise temperature. It looks like the authors are not too familiar with typical microwave calibration techniques which is reflected by the lack of citation of microwave radiometer basics and studies. In operational receivers many strategies for that exist (Maschwitz et al., 2013) as typically gain needs to be adjusted more frequently than TR, relative/absolute calibration.

In the revised version we only use the term "calibration parameters". As mentioned in some answers above, the calibration measurements are performed as frequently as possible during deployments.

P11lEquation: Why so complicated TrˆCCh(C_hot) and not simply Tr – explain the meaning of the different indices.

T_r^{CCh} is derived using different measurements than T_r^{ND}, since they are calculated using different calibration methods, which is shown by using the indices. It is now explained in the text.

P11l19: Give values to underline the statement

This comment refers to the statement that when applying the calibration method that is based on the hot target and the noise diode "[…] two reference temperatures are used, which are above the expected measurement range.

5 As suggested by Reviewer two, we have added a sentence in the beginning of the document, that the atmospheric temperatures surrounding the aircraft (and therefore measured by the MTP), are within a range of 190K - 260K in flight. Higher temperatures – up to 300 K are also possible at very low flight levels. As explained in the instrument description, the hot target has a temperature control keeping its back side at a temperature of 45C (just below 320K), and the noise diode signal is added to this temperature, which is mentioned in this very

10 paragraph. Hence, with the addition of expected atmospheric temperatures, it should now be clear that both temperatures used for calibration are (well) above the expected atmospheric temperatures measured by the MTP. A sentence is added in the revised text, to underline the statement.

P12l4: The calibration strategies might serve different purposes. That the first strategy leads to comparable
15 *results seems astonishing.*
Finding that laboratory values can be used to calibrate flight data is, indeed, astonishing. However, the laboratory data we refer to were produced in very specific conditions, meant to imitate flight conditions. Those results cannot be achieved by a single calibration on the ground, since the trends in changing calibration parameters cannot be reproduced without the specific settings used in our laboratory set-up. It was the purpose of those specific
20 settings to mimic flight conditions as well as possible, and the results show that of all changing parameters, that can influence the instrument during flight, the temperature has the most important influence. This is in agreement with the finding of McGrath and Hewison, 2001, who also found the largest dependence of instrument parameters on temperature changes.

25 *P12l12: The cause for the standing waves is the refractive index of the LN2 – here Küchler et al., 2016 should be cited for details. Here it sounds that just the evaporation is the reason*
We have revised the sentence, so that it now should be clear that the _changing_ interference with the standing wave is caused by the evaporation of the liquid nitrogen. A reference to Section 4.1.1 in Küchler et al., 2016 has also been added, since it gives useful background to the standing wave problem.
30

P12l25: Of course the calibration parameters change with changing environmental conditions if the temperature stabilization of the instrument is not perfect. The question to ask if this is repeatable. Would the same parameters be measured if the instrument had been moved and electrically disconnected in between?
Single calibration measurements, using the hot-cold method were performed before and after a number of
35 campaign deployments, but did not show any consistent picture, due to the large influence of the surrounding temperatures on the MTP instrument. These surroundings cannot be influenced while the MTP is mounted on the aircraft, hence, it is not easy to repeat the measurements needed to establish this consistence. However, all calibration parameters derived during those hot-cold calibration during campaigns were within the range of observed calibration parameters during the cold-chamber measurement series, which is a strong indication that
40 un-mounting the instrument and installing it in the cold-chamber did not have significant impact on the parameters. This is a very strong indication that repeatability is given between different MTP campaign deployments.

P13l29: Why is the temperature unknown – more discussion is needed – see Mc Grawth and Hewison, 2001.
45 We have mentioned the temperature gradient between the back of the target (which is heated), and the (not heated) front of the absorber, to which the measurement is most sensitive to. We have re-worded the paragraph, so that this fact becomes clearer.

P16l8: why do you explain this only here and not at the beginning of the calibration section
The order of the section of uncertainty estimation has been changed accordingly in the revised manuscript.

5   *P16l13: Nobody remembers counts better give the atmospheric temperatures and notate the counts with c_min and c_mac or later c_ref instead of 18500.*
The notation follows the calculation. For better understandability, we have added the corresponding temperatures.

10 *P16l24: "The vertical, grey shaded.." this is not paper style. The figure should be only a reference for the text.*
We have moved the descriptive part of the text to the figure caption.

P17l9: "In literature" then give a reference
The reference to Ulaby, Moore, and Fung, 1981, and Woodhouse, 2005 were already given after the equation
15 was stated. We have moved this reference to the beginning of the sentence.

P17l9 to 29: This paragraph shows that the authors have not much experience with microwave radiometry. It is weird to present the well established radiometer formula at the end and not in the beginning. The formula describes the internal noise of an ideal radiometer and typically one just writes a proportionality and not an equal
20 *sign as other losses occur (e.g. factor 2 for Dicke switching). Further, the authors put in 400 MHz as bandwidth but the double sideband receiver only has 200 MHz in the IF. The most important think to look at the radiometer formula is to check if the noise decreases with longer integration time which is basically what the Allan Variance technique does – it finds out at which point gain fluctuations dominate. This should be checked by the laboratory measurements in the beginning and not in this section. Note, it is strange to only now to provide the integration*
25 *time for atmospheric measurements.*

We have switched the order of the paragraphs, so that it now starts with the theoretical formula. It is correct, that only the 200 MHz bandwith is to be used, which has been corrected. The new uncertainty value is 0.117 K; assuming an atmospheric temperature of 250 (0.108 K for T_atmo = 190 K, and 0.1255 K for T_atmo = 300 K).
30 This is still considerably smaller than the error derived from calibration parameters, so the overall message remains true.
The integration time used for recording measurements was added in the table.

P18l14: If the dominant uncertainty is the noise couldn't it be reduced by longer integration times?
35 There are two possible ways to reduce the noise: 1) to change parts of the instrument, which is not possible due to the aircraft certification. The other possibility is to increase the integration time while recording the signal. This implies a longer recording time for a single measurement cycle, which in turn reduces the horizontal resolution of the measurement, due to the high speed of the jet-engine aircraft the DLR-MTP is flown on. The current settings are recommended by the inventors of the instrument, who have already considered the best compromise
40 between instrument noise figure and measurement resolution (private communication). We have added a sentence to remind the reader of the fact that noise reduction is only available at the expense of horizontal resolution.

P18l30: LO frequency
45

"frequency" was added in the text.

Table 1 does not include all instrument characteristics of interest, e.g. receiver noise temperatures, integration time, polarization. I am missing information on microwave window transmission

The table has been updated.

5

Fig. 8 could be combined with Fig. 10

We have combined the two figures in the revised version.

10 *Fig 11: Different calibrations need to be explained in figure caption. Caption does not say how the difference is calculated (what is the reference – the overall mean?). As I do not see significant temporal development mean and standard deviation could be just added as last lines in Table 6.*

The reference of the shown brightness temperature difference is the time series of brightness temperatures
15 derived with method TTS1. We have updated the figure caption to include a reference to Table 6, which explains the calibration methods.

Reply to the referee #2 comments

We like to thank the reviewer for providing helpful advice to improve the quality of our study. In the following we give our reply regarding the points raised by the reviewer #2. The statements and comments given by the referee are printed in black italics and our comments are presented in blue.

5

Answers to referee #2

General Comments

The paper sometimes reads more like a technical report than a journal article. I would suggest the authors begin with a broader view of such instruments, including their basic operating principles and their scientific applications.

10 *Reference to similar instruments should be included here as well. Then state the motivation for this work and how it supports research with MTP data.*

The intent of our study is to provide the specific characteristics of this particular design of MTP instrument, purchased from JPL. We would like to remind the reviewer that this instrument was not designed by the authors of this study, and this publication is not intended to be a general introduction to the instrument. Such an

15 introduction was already given in the studies mentioned in our publication (e.g. Mahoney and Denning, 2009; Lim et al., 2013), and as a result, we do not see the necessity to include the suggested overview of microwave instruments in use.

The authors note that the MTP was developed by a team at JPL. While the developers have not published comprehensive instrument characteristics, one wonders whether they may have performed some of the work

20 *described in this paper. Have the authors reached out to the developers to understand whether this information exists within the JPL group, and if their results are consistent with the DLR team's findings?*

Indeed, we did have contact to the group at JPL. The instrument design and the components used in this design, differ from those used in earlier designs of the instrument. Earlier versions of the instrument are mounted inside the aircraft cabin, and hence, the investigation of changing surrounding temperatures has not been performed by

25 JPL. Some results and observed instrument behaviour and characteristics were partly discussed with JPL staff, and regarded as expectable. They also compare well to figures and characteristics shown in older documentation, (e.g. the patent for the horn antenna and rotating mirror design from 1976; Fletcher et al.; United States Patent for a Highly Efficient Antenna System Using a Corrugated Horn and Scanning Hyperbolic Reflector; US Patent No. 3,949,404).and the considerations about the instrument design; given in the internal documents/

30 private communication with JPL provided with the transfer of the instrument to DLR.

While interesting, the work presented in Section 5 on sensitivity of LO frequencies and elevation angles seems to be outside the central theme of the paper. After presenting results on performance of various components, calibration methods, and associated uncertainties, it would seem more natural to discuss how performance and

uncertainty impact the final measurement and applications. There is some reference to use of the data for gravity waves and the requisite accuracy for that application, but a more general discussion would make the paper more broadly relevant to readers.

We agree with both reviewers, that Section 5 of the paper fits better as an appendix, as it only presents a very brief investigation of fast-to-apply improvements to the measurement strategy. However, as noted by reviewer 1, there is a lot of room for interesting investigations into implications for radiative transfer calculations and retrieval error estimation, which, if attempted, should be presented in a study of its own. We decided to move this section to an appendix. The in-depth assessment of measurement strategy impacts on the retrieved temperature profiles, including the general discussion about consequences for data analysis, should be considered in a study of its own.

Substantial improvement to the readability of the paper is needed. As noted in Comment 1 above, much of the information is presented as if this were a technical report. Following the Introduction, each section needs to begin with an overview of its contents, motivation for including that content, and how the content fits into the overall purpose of the paper. The material within a section is often not well-organized, paragraphs seem short and choppy, and transitions between topics are lacking.

We have revised the manuscript according to the reviewer's recommendation.

Specific Comments and Questions

P4 eq. 2.1 - Is T the physical temperature? BT is defined here as brightness temperature, but elsewhere in the paper, TB is used (e.g., eq 3.3 on p9).

Yes, this is the Planck equation, based on absolute temperature T.

Following a comment from reviewer #1, we have changed all instances of "BT" in the text to "TB" to be more consistent in the use of the abbreviation of brightness temperature.

p4 line 28 - You state that the target is heated to a constant temperature of approximately 40C. In Table 1 the value is given as 41C. Why not just use 41C in both places?

Looking back at our data, we have corrected both values to 45C, which better represents the temperature of the thermostats used to heat the target (recorded in the MTP's housekeeping data, and shown as orange line in Fig. 8 of the revised manuscript).

p5 lines 1 - The explanation of brightness temperature is awkward and confusing. How about "...which is the temperature of an ideal blackbody emitting the equivalent radiance..."

We have changed the wording according to the reviewer's recommendation Thank you for this helpful suggestion.

p6 line 12 - Reference is made to the antenna diagram. It would be good to direct readers to the corresponding figure (Fig 2, I believe).

We have added the reference to figure 2, both for the antenna diagram, as well as for the earlier mentioned instrument transmission function.

p6 line 13 - "half-sphere" should read "hemisphere"

We have corrected the wording of the revised text.

p8 line 5-10 - It would be informative to share the range of ambient temperatures ex- perienced outside the pod in flight.

We agree. The information (190 – 260K, or warmer, if lower flight levels are flown – up to 300K at the surface) is added in the revised text.

p9 line 6 - "a" should read "at"

The typo is corrected in the revised version of the manuscript.

p10 line 16 - Section 4 includes uncertainty from pointing errors in addition to calibration methods. The title should reflect this, or the point error material should be placed elsewhere.

As suggested earlier, a broader discussion of implications for brightness temperature error is desirable, and some other characteristics influencing the measurement accuracy could also be discussed, as pointed out by Reviewer one. Hence, we have added a new Section (Section 5 in the revised manuscript) to discuss further sources of measurement uncertainty, in which the pointing error is included, as well as some discussion of other error sources, such as synthesizer errors.

p14 line 23 - The sentence that begins with "Note that this definition of usable legs..." is confusing. I'm not sure what you mean.

The chosen flight segments are cut so that no flight manoeuvres lead to large changes in the measured signals. The cutting criteria are based on aircraft parameters only – not on any readings of MTP housekeeping data (as in the laboratory measurements). Hence, times, in which the instrument is still adjusting to new surrounding conditions, are still part of the data used in the investigation. We have changed the wording to make this clearer.

p17 line 9 - This sentence lacks a verb.

The sentence has been corrected.

p22 line 17 - If the authors choose to keep Section 5 as a discussion of new measurement strategy, it would be interesting to demonstrate the impact of LO shifts and/or elevation angle changes on simulated data.

Since the authors decided to skip this section, and to only note the already shown considerations in an appendix, such work should be part of a stand-alone study to assess the impacts of measurement errors and possible changes in the measurement strategy on the retrieval output.

p22 line 31 - "full-with-half-maximum" should read "full-width-half-maximum"

The typo is corrected in the revised version of the document.

p34 Figure 5 - A legend is needed here

A legend with 6 different lines and their explanations would fill quite a large portion of the (already quite busy) plot. Hence, the meaning of the colours and line styles are explained in the figure caption.

p40 Figure 11 and 12 - These figures are too small to differentiate the individual lines/methods.

The two figures are now combined. In the upper panel (Figure 11 in the original manuscript) lines are clearly separated, displaying the offsets between individual methods. In the lower portion of the new figure (formerly Figure 12), the offset-correction is applied, and the plot demonstrates that this is a powerful correction, leading to very similar brightness temperatures being derived from all methods, so that the lines partly overlap.

p44 Figure 15 (left panel) - It's impossible to distinguish the 58.363 GHz line from the 56.363 GHz line

The 58.363 GHz lines are plotted in the background, and similar to lines of the other frequencies – hence, they are not well visible. We have added a sentence in the figure caption to point this out to the reader.

p45 Figure 16 - The legend indicates lines for 6 altitudes are shown, but I can only see 4 on the left plot.

At horizontal viewing angle, there is such a small difference between the lines representing altitudes at or below 8 km overlap in the left panel. We have added a sentence in the figure caption to point this out to the reader.

15

II: List of changes in the Manuscript

The reviewers have suggested a number of changes to the manuscript, many of which are mentioned in our answers to the reviewer comments, above. Due to the substantial text work on the revised manuscript, we would like to only give an overview the most important changes:

1) The **Abstract** is slightly re-worded.
2) In the **Introduction** we now give a brief context of the use of the MTP instrument on HALO. Other parts of the Introduction have been slightly shortened to maintain the overall length.
3) Section 2 (**Instrument description**) was rearranged, and the concept of measurements is now better explained. Following the reviewers' recommendations, a brief overview is added at the beginning of the section to introduce what is being shown, and how it fits into the scope of the paper.
4) Also, in Section 3 (**Characteristics of the wing-canister MTP flown on HALO**) a brief overview is added at the beginning of the section to introduce what is being shown, and how it fits into the scope of the paper. The discussion of instrument noise characteristics was re-written, and now contains a few comments on the use of the Allan Variance.
5) The introduction of Section 4 (**Investigation of calibration methods for the HALO-MTP**) was rewritten to better introduce the pros and cons of the different approaches to calibrate MTP data. The sub-sections are rearranged in a more easy-to-follow logic, and remarks about thermal stability of the heated target are added in the respective section, following the reviewers' recommendation.
6) The **discussion of uncertainties** resulting from calibration now starts with introducing the theoretical values, before exploring the calibration methods. The discussion of other error sources is now moved into a new Section (see below)
7) Following the reviewers' recommendations, Section 5 (**Range of sensitivity**) has been moved to an appendix.
8) A **new Section** (Further considerations relevant for retrieval set-ups) was written, which now contains a discussion of uncertainties resulting from other sources, except calibration. Is has three subsections, discussing pointing error, synthesizer errors, and measurement settings impacting the retrieval output (i.e. a brief overview of the main results from the Appendix).
9) The **Summary** section was shortened.
10) The **acknowledgements** are extended to properly acknowledge the support from the MTP groups at JPL and NCAR, who have given advice in the operation of the instrument.
11) Some new **references** are added in the text (as well as in the reference list), following the reviewers' recommendations (see answers to the reviewer comments, above)
12) Figures 8 and 10 of the original manuscript are now **merged into a single figure** (Figure 8 in the revised manuscript)
13) Figures 11 and 12 of the original manuscript are now **merged into a single figure** (Figure 10 in the revised manuscript)

III: Marked-up Manuscript

Measurement Characteristics of an ~~airborne~~ **Airborne** Microwave Temperature Profiler (MTP)

5 Mareike Kenntner[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 (Mareike.Kenntner@DLR.de)

10 **Abstract.** The Microwave Temperature Profiler (MTP), an airborne passive microwave radiometer, ~~records~~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) ~~assessed.~~ 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.

~~For this instrument a thorough analysis of instrument characteristics has been made. This is necessary to correctly determine the accuracy and precision of MTP measurements, and crucial for a retrieval algorithm to derive vertical profiles of absolute atmospheric temperatures.~~

20 -Using a laboratory set-up, the frequency response function and antenna diagram of the instrument was 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. ~~Furthermore, using the radiative transfer model Py4CAtS, the sensitivity to the atmospheric layers around flight altitude was investigated.~~

25 ~~It was found that using the standard measurement settings, the DLR MTP's vertical range of sensitivity is limited to 3km around flight altitude, but can be significantly increased by adjusting the standard measurement strategy, including slightly weaker oxygen absorption lines and a different set of viewing angles. Calibration parameters do clearly depend on the state of the instrument~~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.

~~With here~~ By 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 ~~atmospheric temperature around flight level~~

with a precision of 0.38 Kbrightness 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, using the radiative transfer model Py4CAtS in a pencil-beam approach, indicate that the vertical observation range of the MTP instrument can be increased by applying a modified measurement strategy.

5   This is the first time such a thorough instrumentan extensive characterisation of a MTP instrument, including a thorough calibration strategy assessment, is published. With theThe presented results, itrelevant for the wing-canister design of the MTP instrument, are important when processing MTP data: Knowledge of the relevant uncertainties and instrument characteristics is now possibleessential for retrieval setup and mandatory to correctly identify and interpret significant atmospheric temperature fluctuation signals in MTP data and choose the best possible measurement strategy fitting the

10   purpose of the measurement campaign..

1 Introduction

Aircraft campaigns have long been used to study atmospheric composition and dynamics. 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. One such instrument is the passive microwave temperature profiler (MTP; Denning et al. 1989), which provides temperature profile information both, above and below flight level. Due to its measurement principle, it is largely independent from the prevailing atmospheric conditions, such as clouds or sunlight, which hardly interfere with the measurement – an advantage of the MTP in comparison with other remote sensing techniques, such as LIDAR instruments. Such an instrument has been purchased and deployed by the German Aerospace

20   Center (DLR).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.

25   1989) complements other instruments such as the Basis 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

30   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 differential absorption lidar (WALES; Wirth et al., 2009) offer 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 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, 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. ~~The MTP records thermal radiation emitted by oxygen molecules in the atmosphere. Its level 1 data output is a brightness temperature, which has to be converted to absolute temperature profiles by using a retrieval algorithm, that utilises forward radiative transfer calculations. For correct interpretation of the structures found in the retrieved temperature fields, it is necessary to have precise knowledge of the instrument characteristics, such as the instrument response function, and the precision and accuracy of the brightness temperature measurements that are the input to the retrieval algorithm.~~

~~MTP data have been used for a variety of purposes, both directly interpreting MTP data or using it to support studies of data from other instruments or model output, as the following brief overview will show.~~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). ~~Studies~~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~~). After the Airborne Antarctic Ozone Experiment (AAOE), Hartmann et al. (1989) calculated potential vorticity on potential temperature isentropes from MTP measurements combined with pressure and wind measurements provided by the meteorological measurement system of the aircraft to gain insight in mixing processes within the polar vortex. Davis et al. (2014) investigated~~), 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.

Furthermore, MTP measurements have been utilised to investigate gravity waves ~~(GWs)~~ 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 ~~performed and compared to cases in which MTP measurements are available. These studies~~published, aimed at improving the understanding and numerical description of ~~GWs. Such~~atmospheric gravity waves,

including studies have been undertaken 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 these studies focusing on mesoscale temperature fluctuations or vertical temperature gradients, precise knowledge of the instrument characteristics, such as intrinsic noise, and 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 consistentcontinuous use of data from various MTP instruments in many studies over the past decades, a thorough instrument characterisation and derivation of measurement accuracy has not yet been published. For the first time, this study presents all those relevant instrument characteristics for an of the HALO-MTP instrument. ThisAll 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 DLR-MTP. This includesHALO-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), discussinga 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. Finally, radiative transfer calculations are used to determine whether the current measurement settings are already ideal, discussing possible improvements (Section 5). Furthermore, a brief description of the instrument and its measuring principle will be given in Section 2.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. Two such MTP instruments have been built, and one of them has been acquired by the German Aerospace Center (DLR).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 emitted by oxygen molecules in the atmosphere. As ~~all~~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 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 (BT~~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. 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).

~~2.1 Basic MTP instrument components~~

~~Table 1 lists all important parts of~~In the MTP instrument ~~used to record the MTP data during flight. A,~~ a horn antenna is used as the receiver for the incoming atmospheric radiation. ~~Through down conversion~~The measurements are based on the heterodyne principle, which means that through mixing with a defined frequency, the local oscillator frequency (LO), ~~and~~the incoming signal is converted to a base-band. Then, by low pass filtering only the part of the incoming radiation spectrum around the current LO is ~~let through.~~selected. The MTP uses a double-sideband receiver. Thus, frequencies below and above

the LO are down-converted to the same base-band, so that both flanks of an oxygen absorption line can be further processed. The passing signal is converted to a voltage, which is proportional to the squared input ~~intensity~~amplitude, representing the power. This voltage is finally translated to a digital count number, stored in the MTP data file, and later ~~translated~~converted into a brightness temperature through calibration (see Section 4).

5 ~~Using a rotating mirror in front of the instrument's antenna (number 3 in right panel of Figure 1), the direction from which the radiation is collected can be changed, making the derivation of altitude resolved temperature profiles possible. In its standard deployment settings, as programmed in the JPL instrument software, ten viewing angles are being used during one measurement cycle; five above the horizon, four underneath, and one pointing exactly towards the horizon. At each angle, measurements at three LOs, corresponding to the frequencies of three strong oxygen absorption lines, are made (see Table~~
10 ~~1), before moving to the next elevation.~~

The temperatures of the important radiometric parts of the radiometer, such as the mixer, synthesizer, 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).

15

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
20 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.

25

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 ~~40~~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
30 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 ~~BT, which is the temperature, an ideal black body would have, which emits the measured radiance, according to Eq. 2.1.~~

~~This BT~~TB, which is usually not equal to the outside air temperature, since the measured signal is influenced by multiple ~~altitude~~layers of the atmosphere. To derive absolute temperature from the radiation measurement, ~~forward~~radiative transfer

24

calculations have to be carried out and compared to the measured radiances~~. This happens in~~ by applying a retrieval algorithm ~~that is applied~~ 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. ~~Retrieval methods~~However, the discussion of retrieval algorithms and related uncertainties ~~will not be discussed~~are beyond the scope of this paper.

~~2.2 Wing~~ Specific wing-canister instrument hardware characteristics

~~The design of the DLR-MTP was first introduced in 2008 and allows mounting the MTP inside a canister underneath the wing of a research aircraft. Details of the instrument design can be found in Mahoney and Denning (2009). The wing-canister instrument is not the first design of the MTP.~~ Differences to older instrument designs are ~~mentioned~~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 standard LOs and elevation angles used in a measurement cycle are summarised in Table 1. Two instruments were built using this new design. 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 DLR MTP) was acquired by DLR, and has been flown on the German research aircraft HALO (Krautstrunk and Giez 2012).~~

The filter ~~band with~~bandwidth of the ~~DLR~~HALO-MTP is fixed to ±200 MHz around the LO, with a gap of approx. 20 MHz at the line ~~center. However, the radiometer architecture using a mixer to down shift the incoming signal allows measurements at various frequencies, depending on the chosen LO.~~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 ~~signal~~ is doubled twice, 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.

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 ~~DLR~~HALO-MTP can ~~run~~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 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 functions 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 change the relation between the recorded signal and the source temperature, despite the fundamental assumption in MTP calibration that this relation is always linear. Moreover, the noise diode used for calibration may change its signal (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.

~~The~~ (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

5 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 ~~to have power so~~such that the MTP signal was well ~~over the~~above its inherent noise level. For the measurement of the filter function, the source frequency was tuned between

10 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 Friis Transmission Equation (e.g. Balanis ~~(~~ 1997):

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

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

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

15 The resulting instrument transmission functions for the three standard ~~LOs~~frequency channels are shown in Figure 2, left panel. It shows symmetrical shapes for all ~~LOs~~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. 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

20 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

25 half maximum (FWHM; red, dashed lines in Fig. 2 middle and right panel) of the antenna diagrams. It is actually mainly defined by the 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

Changing surrounding temperatures of the MTP can influence the performance of the instrument: Amplifiers may change the relation between recorded signal and source temperature, despite the fundamental assumption in MTP calibration is that this relation is always linear. Moreover, the noise diode used for calibration may change its signal, and the overall instrument noise can be affected.

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 onea 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'.

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. This adjustment can take up to 15 minutes after the initial temperature change. To ensure that results from the laboratory measurments are representative for instrument deployment on an aircraft, they will be compared to an analysis of mission data from the ML CIRRUS 2014 campaign.

3.2.1 Linearity of the sensor

Using the measurements of the two ambient targets installed within the chamber, it can be shown that for the ~~DLR~~HALO-MTP the linear relation between the source temperature and the measurement output is given at all standard ~~LOs~~frequency channels (see Fig. 5). Since not only the temperature of the target changed during this test, but also the temperature of the sensor unit itself (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 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 ~~a~~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 ~~representing the instrument state.~~. 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.~~2.2~~3 Noise characterisation

~~Using the same time series of cold chamber measurements the~~The instrument's noise figure was characterised~~,~~ using the signal measured when pointing towards the hot target. ~~When pointing towards a calibration target at a stable~~ It is assumed that due to the temperature stabilisation of the target, the mean measurement signal should not change over time~~, and~~. 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. ~~Slight changes in system temperature over time have to be~~, 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). ~~The~~Using all segments of the cold-chamber measurements, the resulting ~~DLR~~HALO-MTP noise figure, as shown in

5 Figure 7 (top), can be characterised by a Gaussian distribution with a standard deviation of approximately 6 cnts and the mean at 0 cnts.

The same method as for the cold chamber measurements is used for ~~DLR~~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

10 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

15 middle panel of Figure 7 shows the noise characteristics at ~~LO~~56.363 GHz. ~~Plotted is also a Gaussian function with a mean at 0 cnts, and a standard deviation of 6 cnts, as implied by the cold chamber noise figure (green line).~~ 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 ~~LO.~~the frequency channel. This is strong evidence that the ~~DLR~~HALO-MTP noise figure does not change between flights, and the laboratory characterisation can be used to determine long-term

20 stability of the instrument in between campaigns.

For the spectral analysis of the 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 ~~LO~~56.363 GHz ~~of the DLR MTP~~, as shown in Fig. 7 (bottom), reveals that the

25 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~~P_k, is given by:

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

For the three standard ~~LOs~~frequency channels, the lag-1 autocorrelation of MTP measurements during the ML CIRRUS campaign is $\alpha \cong 0.7$.

30 Fit parameters characterising the noise figure at the three standard ~~LOs~~frequency channels are summarised in Table 2.

With the above findings, characterising the ~~DLR~~HALO-MTP 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

5 deviation of the Gaussian distribution of noise values can be used to determine the variance of ~~BTs derived from the raw signals, once the calibration parameters are known~~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

10 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 ~~DLR~~HALO-MTP

15 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 ~~line~~calibration parameters also bears the danger of inconsistencies under rapidly changing measurement

20 conditions, which could lead to large errors in the calculated ~~BTs~~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 ~~are being tested.~~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:

25

~~For the DLR-MTP there are three possible calibration strategies that can be used to determine the line parameters:~~

    ~~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~~, and using the equations:~~

30 ~~$T_B^{CCh}(c)$~~$= s_{\text{cal}}^{CCh}(c_{\text{hot}}) \cdot c - T_R^{CCh}(c_{\text{hot}})$ $(Eq\ 4.1a)$

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

31

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

2.1. ~~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 laboratory measurements within a similar instrument state.~~ This is the standard calibration method of radiometers in a stable environment. Using this method to calibrate the sensor, before ~~making~~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 ~~Furthermore, this~~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 ~~will be~~ shown below. However, this method can be used to check the overall health of the instrument in between deployments.

3.2. 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}

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

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

~~Using a noise diode to add a measurement signal, representing a known temperature difference to the temperature of the hot target,~~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.

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

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

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

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.

TheOther calibration methods, such as the tipping curve calibration (e.g. Küchler et al., 2016) are not available for the DLR MTP due to the given instrument design, and the need for an efficient measurement strategy.

In the following sections will show that each , the various methods 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 mentioned calibration strategies leads to comparable results, but differ in their representation of the instrument state or the determination of uncertainty.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 at the end of the section will summariseof uncertainties and summary of the results, leading to an assessment of a best practice for calibration of DLRHALO-MTP data after a campaign. This permits the best possible representation of the instrument state during flight, leading to the best possible measurement accuracy and estimation of uncertainty, before applying a retrieval algorithm to derive absolute temperatures., 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 DLRHALO-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:

33

(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 cyclemeasurement 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 do indeed change with the scanning unit temperature. That corroborates the assumption, that DLRHALO-MTP flight data cannot simply be calibrated by using fixed calibration line 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. Whether this is sufficient to represent the changing conditions during flight has to be tested using mission data (see Section 4.3).

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. The first is that i) the ND offset signal is the same each time the calibration measurements are performed. The second assumption is that and ii) the BTTB measured when pointing towards the heated target corresponds to the measurements of the temperature sensors at the back of the target.

If one of those assumptions is incorrect, large calibration errors can occur due to the fact that the two points used to determine the calibration line parameters are both at quite high temperatures: 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.

Both assumptions are tested in the following, using the calibration measurements performed in the cold-chamber.

4.2.1

Noise diode offset temperature

The ND offset signal has to beis characterised using the hot-cold calibration method. To investigate the influence of changing conditions used during a measurement flight, the cold chamber measurement series is used, 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 DLRHALO-MTP ND offset temperatures are shown in Fig. 9. For better comparability, the means of the temperature and count values have been removed. Those

~~correspond to the reference values in Table 4 (column 2 and 3).~~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).

5   Again, it is possible to apply a linear fit between the recorded ND offset signal, $\hat{c}_{ND} = c_{ND} - c_{hot}$, 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 for any of the three ~~LOs.~~frequency channels. This spread translates into the remaining uncertainty in the ND offset temperature.

10   ~~4.2.2~~

~~Hot target temperature measurement~~

~~The housekeeping data of the MTP indicate large temperature differences between the air in front of the hot target and the heated back. This difference can reach up to 55K, so that 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~~
15   ~~front of the absorber, of which the exact temperature is unknown. To investigate the hot target measurement characteristics the calibration parameters, determined from the hot cold calibration method, are used to calculate the hot target BT associated with the current measurement signal.~~

~~Indeed, Figure 10 shows the clear trend towards colder BTs with lower scanning unit temperatures, which correspond to a colder environment of the MTP instrument. This is contrary to the readings of two Pt100 temperature sensors, which shows~~
20   ~~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 10). The difference between the readings of the Pt100 sensors in the rear of the target and the correct BTs measured during calibration can be as large as 3 K. Still, the linearity of the sensor again allows for a linear fit between the current scanning unit temperature and the average associated hot target BT. Thus, in flight calibration can be performed, using a corrected hot target BT, according to the MTP instrument's housekeeping data. The parameters to correct~~
25   ~~the hot target BTs used in the calibration are shown in Table 5.~~

~~4.3~~ 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 ~~is~~can be used instead. This temperature is interpreted as the ~~BT~~TB measured at the 0° elevation (horizontal measurement). ~~In Section 5 it will be~~
30   ~~shown~~Simple radiative transfer calculations show, that the MTP measurements at all standard ~~LOs~~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 ~~using~~combining the calibration measurement while pointing at the hot target~~, combined~~ 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.~~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 ~~were~~are compared to find the best practice of deriving ~~BTs~~TBs from MTP raw counts~~. All~~ by applying all eight methods ~~have been applied~~ to the same set of mission data~~, using~~. 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 ~~between cycles~~during a measurement cycle, and no curves were flown (roll smaller than 5°). Note that this definition of usable ~~legs~~segments is not based on any parameters connected to the ~~DLR~~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 ~~of~~with at least 10 minutes ~~length~~duration (i.e. including at least 50 measurement cycles) could be identified. The ~~BTs~~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 ~~BTs~~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. ~~11. For plotting purposes, the difference between the BTs derived with each individual calibration method to the HALO TS is shown. The BTs~~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. ~~This leads to~~

~~the assumption~~, indicating that differences in the respective calibration coefficients affect the accuracy of the derived ~~BTs~~TBs more than the precision.

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

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

with $T_{\mathrm{B}}(\nu_{LO}, \alpha)$ and $T_{B}^{corr}(\nu_{LO}, \alpha)$ denoting the original and the corrected ~~BTs~~TBs under elevation angle α and at a specific ~~LO,~~frequency channel (i.e. LO), respectively. $\overline{T_{\mathrm{B}}(\nu_{LO}, 0°)}$ denotes the leg-mean of the original ~~BTs~~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 ~~BTs~~TBs will still contain individual small-scale structures, which might differ from those in the HALO TS measurements. ~~As a result, this correction will not influence further analysis of mesoscale structures.~~ For individual calibration strategies, the subtracted offset can be as small as 0.8 K or as large as 7 K (see Fig. ~~11~~10, top panel). The good agreement between all eight corrected ~~BTs~~TBs under the different viewing angles (see Fig. ~~12~~10, lower part), indicates that removing the offset will not significantly change the shape of the temperature profile calculated in the retrieval. Moreover, the accuracy of temperature measurements now matches that of the HALO TS, which has an overall uncertainty of 0.5 K (Ungermann et al., 2015).

~~With~~

After applying the offset correction ~~applied, plotting,~~ the RMS difference between the 0° ~~BTs~~TBs and HALO TS, as shown in Fig. ~~13~~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, ~~to~~ 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 ~~BTs~~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° ~~BT~~TB and the leg-mean HALO TS. The deviation between the resulting 0° elevation ~~BTs~~TBs and HALO TS is ≤ 0.38 K at all three ~~LOs~~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.~~5~~4 Discussion of uncertainties and measurement precision

~~To estimate the uncertainty of the calibrated BTs used as input to the retrieval algorithm, a number of contributions have to be taken into account. Some result from calibration parameter derivation, others are inherent to the instrument itself. Both will be discussed in the following.~~

~~**4.5.1 Uncertainty resulting from calibration parameters**~~

~~Uncertainties arise from the use of the different reference temperatures used in the calibration process and are 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 changing instrument states. Hence, a quadratic sum of the individual errors is not suitable and will lead to a large over-estimation of the total BT error. Hence, a sensitivity analysis to estimate the overall 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, BTs 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.~~

~~Two control calculations are made, adding the corresponding 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.14), or as flat as possible (s_{cal}^{min}, blue lines in Fig. 14), following:~~

$$s_{cal}^{max} = \frac{(T_1 + \Delta T_1) - (T_2 - \Delta T_2)}{(c_1 - c_2) - 2\Delta c} \quad (Eq\ 4.\underline{5}~~a~~)$$

$$T_R^{max} = (T_2 - \Delta T_2) - s_{cal}^{max} \cdot (c_s + \Delta c) \quad (Eq\ 4.\underline{5}~~b~~)$$

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

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

~~assuming that T_1 (with associated measurement signal c_1) is the warmer temperature used in the calibration. Comparing the BTs of the reference calculation to those of the two control calculations reveals the maximum uncertainty in the derived BTs. Furthermore, in parallel to the offset correction introduced in the previous section, a BT correction for the control calculations is introduced: Here, the offset correction is calculated from the difference between the BTs of the control calculation (T_B^{ctr}), and that of the reference calculation (T_B^{ref}), at 18500 cnts:~~

38

$$T_B^{corr}(c) = T_B^{ctr}(c) - \left(T_B^{ctr}(18500) - T_B^{ref}(18500)\right) \quad (Eq.\,4.~~7~~)$$

~~Results for all three calibration methods are shown in Figure 14. The shading around the two lines resulting from the two control calculations indicate the error range induced by 6 cnts uncertainty of the measurement signal that is to be calibrated (see Sect. 3.2.2 and Table 7). The vertical, grey shaded region indicates the expected range of counts within one measurements cycle, if the horizontal measurement is at 18500 counts. Within this region, the resulting error, indicated by the upper most and lower most edges of the blue or red shaded region, 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 BTs. However, the calibration method relating to the cold chamber measurements is most reliable in the case that the measured signals deviate largely from the measurement signal at the horizontal elevation (i.e. if large vertical temperature gradients are present around the current flight level). The overall uncertainty is clearly below the already established value of ~ 0.38 K for all methods, mainly caused by the measurement noise.~~

In the literature~~, a different approach~~ (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:

$$\sigma_N = \frac{T_{\text{sys}}}{\sqrt{\Delta f \cdot \tau}} = \frac{T_R + T_{\text{atmo}}}{\sqrt{\Delta f \cdot \tau}} \quad (Eq.\,4.~~8~~5)$$

in which Δf denotes the filter bandwidth, and τ represents the integration ~~time (e.g. Ulaby et al., 1981; Woodhouse, 2005).~~. The ~~DLR~~HALO-MTP has an ideal filter width of $\Delta f = $ ~~400~~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 $\sigma_{N,theo} = 0.$~~0827 K, which is approximately four times smaller than the value established through the calibration of mission data. However,~~117 K, However, these values used to derive the theoretical variance 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 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 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:

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

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

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

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

5 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):

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

10 Results for all three calibration methods are shown in Figure 12. Within the typical region of measurements (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

15 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). 2). Moreover, itFor 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, 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

20 also confirms that the uncertainty of derived BTsTBs 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.

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.

5

Uncertainty from pointing error

The position of the DLRHALO-MTP instrument underneath the wing of the aircraft makes it sensitive to the altitude and speed of the aircraft, which alters the pressure underneath the wing, leading. The bending and torsion of the aircraft body parts leads to deviations between the assumed pointing of the instrument and its true viewing direction. On the ground, this

10 deviation offset can easily be determined. During flights, the measurement of the inertial sensor, which is part of the DLRHALO-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

15 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...

4.6 Summary

In the previous sub-sections the influence of changing outside air temperatures on the instrument state was investigated. A

20 series of measurements in a cold chamber was used to simulate in-flight conditions. It was shown that the linear relationship between the source temperature and the measured signal is maintained. Still, 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

25 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 this of the HALO static temperature, and a precision better than 0.38 K. It has further been shown that the dominant source of measurement uncertainty is measurement noise.

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. 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 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 pick up information on the state of view deeper into the atmosphere) is smaller than implied by the retrieval output, which is given at an altitude range of ±8 km around the flight altitude of the aircraft.

The. 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 LOsfrequency 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 ~~of data recorded with the NCAR MTP~~ 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 range of sensitivity depends on two settings in the measurement strategy: The set of LOs at which the measurements are taken, and the set of elevation angles used. The latter influences the vertical resolution of the temperature profile that can be retrieved from the MTP measurements, while the choice of LO influences the altitude range that the measurement is sensitive to. Within radiative transfer (RT) calculations this is determined by the transmission function~~

The choice of frequency channels can have consequences on the calibration options, but can lead to clearly enhanced vertical observation range 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.

~~, $\mathcal{T}(v) = \exp(-\tau(v))$, which describes the ration of outgoing and incoming radiation for a layer of the plane parallel atmosphere. It is expressed through the optical depth $\tau(v)$ defined as the integral of the absorption coefficient (α), which depends on the frequency (v), the length of the path (s') through one layer, the pressure (p) and temperature (T) of the layer within the plane parallel atmosphere (e.g. Schreier et al., 2019):~~

$$\tau(v) = \int_0^s \alpha(v, s', p, T) ds' \quad (Eq.~5.1)$$

~~To investigate the range of sensitivity of the DLR-MTP, it is useful to calculate the signal contribution from each respective layer of the atmosphere, determined by the weighting function (WF) defined as:~~

$$W(v, s) = \frac{\partial \mathcal{T}(v, s)}{\partial s} = \alpha(v, s) \cdot \exp(-\tau(s)) = \alpha(v, s) \cdot \exp\left(-\int_0^s \alpha(s') ds'\right) \quad (Eq.~5.2)$$

~~The WFs for the three standard LOs used by the DLR MTP under the nine non-horizontal viewing angles used in the standard measurement strategy and assuming an aircraft altitude of 11 km are shown in Figure 15 (left panel). For RT calculations needed to derive the WFs, the Python scripts for Computational Atmospheric Spectroscopy (Py4CATS[1]; Schreier et al., 2019) are used. The WFs were computed from absorption coefficients using spectroscopic line parameters from high resolution transmission molecular absorption database (HITRAN; Rothman et al., 1998), assuming a mid-latitude summer atmosphere (Anderson et al., 1986). To make them comparable the WFs are scaled so that the sum of weights for~~

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

each viewing angle and LO equals 1. 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 LOs, information on the vertical temperature profile can still be 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 this distance or further away. At lower altitudes, with higher pressure leading to less transmission beneath the aircraft, this distance is even less.

5.1 Choice of LO frequencies

Logically, the best idea to widen the range of sensitivity would be to use different LOs that are located at weaker absorption lines than the standard LOs, on the edge of the 60 GHz oxygen absorption complex or even between two lines, as was done with the older MTP instruments. The choice of an LO at the center 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 calculations, 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 the atmospheric layers between ground and 110 km altitude were all set to 250 K. The simulation is made, using TIRAMISU (Xu et al. 2016), a retrieval algorithm developed to process MTP brightness temperatures, which uses the radiative transfer model GARLIC (Schreier et al. 2014). 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 LOs 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 a measurement 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 small part of the measured signal.

The resulting brightness temperatures are shown in Figure 16. 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 16 shows that for any LO 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 LOs at

44

frequencies above 54 GHz should be considered. The results for the near-zenith measurements (right panel of Fig. 16) indicate that the atmosphere is partly transparent for all possible LO frequencies 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 LOs 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. 15 (right panel). Obviously, the new LOs 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 LOs. This is due to the partial transparency of the atmosphere at these frequencies, indicated by low BTs in Fig. 16, combined with the fact, that the viewing direction points through a medium that becomes optically thinner with increasing distance to the sensor.

5.2 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):

$$l_{rel80°} = \frac{\cos(10°)}{\cos(90° - \alpha)} \quad (Eq. 5.3)$$

The relative path lengths to the ±80° angle are summarised in Table 8. Especially the three largest elevation angles used in the standard MTP measurement strategy (underlined values in Table 8) 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. 17); 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 8, including one angle with $l_{rel80°} \approx 1.5$. 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°.

45

5.3 Determining a new measurement strategy

Using the findings of the above sections, redundancy in information within one measurement cycle of the DLR-MTP can be reduced. Moreover, at least below the aircraft, the range of sensitivity of the MTP measurements can be significantly enlarged by using at least one LO at the frequency of a weaker oxygen absorption line.

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 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 LOs 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 9. 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 LOs 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. 17. Depicted are the cumulative weights to indicate the percentage of the measurement signal that is acquired with increasing 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. 16, where absolute values are shown.

Strategy '8E4LOa' shows the result of simply adding a LO to the standard set (Fig. 17a). In the other three proposed strategies, only two LOs of the original set are kept, and two LOs 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. 17 d) are to be avoided. Also, if the aircraft is flying at lower altitudes, the LO at the weakest absorption line might be influenced by the surface temperature. Hence, depending on the planned flight pattern, strategies '8E4LOb1' (Fig. 17c)) or '8E4LOc' (Fig. 17b) should be favoured. 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.

6 Summary

This work 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 results are a necessary basis for the analysis and interpretation of temperature profile data retrieved from MTP measurements, which are used to assess the state of the atmosphere around flight level of the research aircraft, and to investigate mesoscale temperature fluctuations in the atmosphere.~~ 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 ~~region on which the instrument can collect information~~vertical observation range, is fundamental to the correct set-up of a retrieval algorithm, as well as for the interpretation of the ~~measured~~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). ~~Both measurements are crucial parts of the retrieval algorithm used to derive absolute temperature profiles from MTP raw data.~~ 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-~~with~~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.

~~Laboratory measurements~~

Measurements in a cold-chamber set-up, as well as data recoded during a field campaign, were used to characterise the MTP's noise figure (Section 3). ~~After removing a drift on the data, caused by the instrument adapting to changing surrounding temperatures, the remaining noise signal~~It can be described by a Gaussian distribution with mean of zero counts, and a standard deviation of 6 counts. 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.~~ The~~, however, the presented characterisation of the ~~DLR~~HALO-MTP noise ~~figure is sufficient to identify~~allows the identification of significant atmospheric signals in MTP measurement time series.

~~The~~ Furthermore, in the laboratory measurements ~~were also used to proof~~ 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 ~~in a series of measurements in a cold chamber~~ (Section 4~~), in which the changing temperatures surrounding the MTP during a mission flight were simulated. An influence on the instrument state was indeed found, so that~~). The measurements revealed clear changes in all calibration parameters ~~change~~, depending on the ~~surrounding conditions. However, linear relationships between the calibration parameters and the~~

47

instrument state (expressed through the readings of a housekeeping ~~cold chamber~~ temperature ~~sensor, or~~ . This includes a change in the measured ~~signal while pointing towards the heated calibration target), could be established, and can be used in the conversion of raw measurement signal to~~ brightness ~~temperatures. Furthermore, a similar correction was established to account for the~~ temperature ~~gradient present in the heated calibration target.~~

5 ~~Three different possibilities to perform the calibration are presented and compared. One of them refers to the cold chamber measurements, and proofs to be extremely reliable~~ when ~~tested with campaign data from the ML CIRRUS campaign. The other two methods utilise the~~ pointing towards the built-in ~~heated~~hot calibration target, ~~and provide~~as well as a ~~more direct relation to the~~change in ~~flight~~instrument state. The receiver noise ~~diode, used in one of those in flight calibration methods, was characterised through hot cold calibration. It was~~ temperature caused by the electrical parts, and the calibration slope,

10 caused by a change in amplification of the signal. Corrections to account for those changes have been found, and it could be shown, that ~~the noise signal depends on the instrument state (surrounding temperature), implying that a correction has to be used when calibrating mission data. Again, a linear relationship was found that provides the relationship between the measured noise diode offset signal and the correct associated~~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

15 temperature ~~offset needed for the calibration.~~

~~The~~, and a precision better than 0.38 K. To achieve this accuracy, the necessity of an offset-correction relative to HALO TS has been identified, both for the calibration relating to laboratory measurements, and ~~that~~using the noise diode signal. The correction procedure was introduced as comparing the leg-mean of the calculated ~~BTs~~TBs at 0° elevation angle (horizontal measurement) to the leg-mean of HALO TS.

20 ~~The offset corrected brightness temperatures were compared to HALO TS using data from all ML CIRRUS mission flights. The RMS difference to the HALO TS is found to be between 0.25 K and 0.37 K. This range is valid for all calibration methods, as long as the ND functions properly. Otherwise the calibration methods using the ND offset signal cannot be applied and will lead to false results. It has been shown that the offset corrected BTs of all calibration methods agree well within this specified RMS range at all elevation angles used during the ML CIRRUS campaign in 2014. Hence all three~~

25 ~~calibration methods produce comparable results.~~

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

30 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. ~~Using this method, BTs can be derived with a precision better than 0.38 K, and an accuracy matching that of the HALO TS measurements, which have an estimated overall error of 0.5 K (Ungermann et al., 2015).~~

When analysing the uncertainty of the calibrated brightness temperatures (Section 4.~~5.1~~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.

5

~~In Section 5, improvements to the measurement strategy were proposed, after showing that the current standard settings do not provide ideal input to the retrieval algorithm used in the processing of MTP data. While limitations to the range of sensitivity have already been indicated in earlier publications related to the MTP, a theoretical study involving radiative transfer calculations to infer the weighting functions for the MTP instrument has not been published so far.~~

10 ~~It could be shown that using the standard settings the weighting functions of all measurements, under all elevation angles and local oscillator frequencies,~~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 $\sim$ 3 km region around flight altitude (i.e. $\pm$1.5 km around flight level).

15 A proposal to improve the measurement strategy for future missions of the MTP ~~has been~~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 weighting functions connected to these new measurement strategies imply~~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-

20 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 instrument parameters and characteristics needed to accurately analyse and interpret the data produced by HALO-MTP measurements. It is the basis to understand measurement uncertainty, the (vertical) range

25 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 ~~with~~to in-situ as well as other remote-sensing measurements made on the same

30 mission flights.

49

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.

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.

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.

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

Leutbecher, M. and Vokert 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 HITRAN MOLECULAR SPECTROSCOPIC DATABASE AND HAWKS (HITRAN ATMOSPHERIC WORKSTATION): 1996 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.

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., Höpfner, M., Jurkat, T., Kaufmann, M., Kaufmann, S., Kleinert, A., 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
5   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.

10   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.
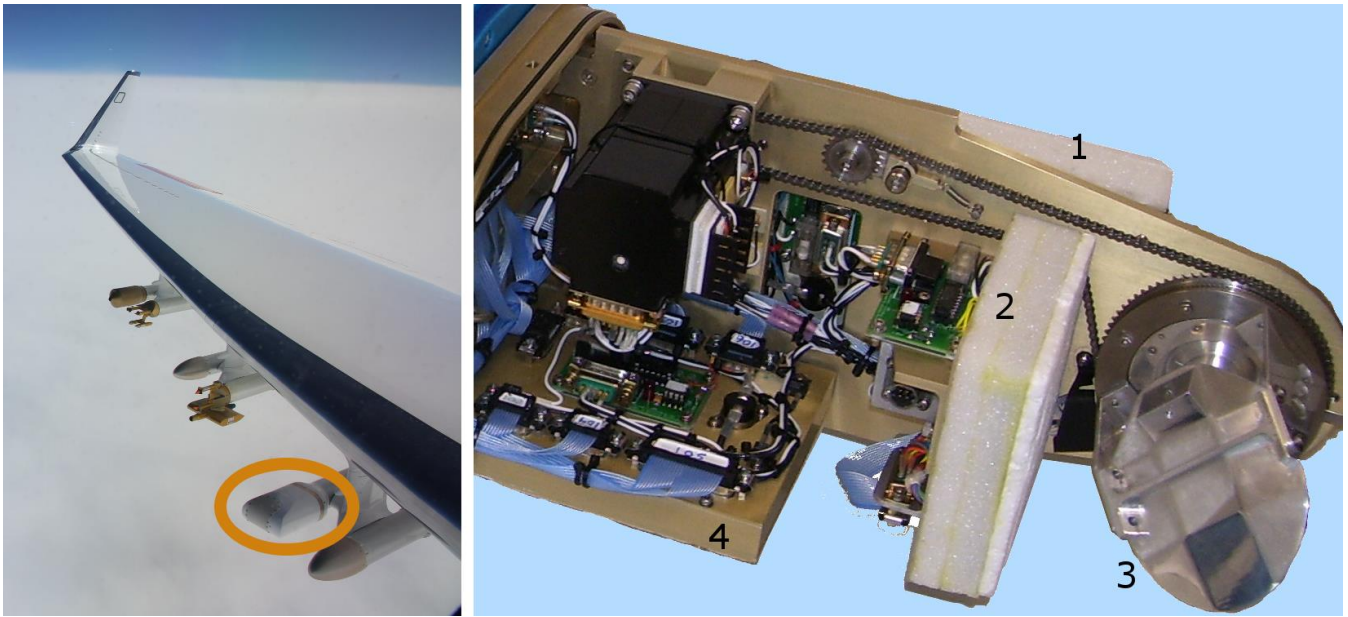
Figure 1: ~~DLR~~HALO-MTP instrument. Left: Position of the MTP underneath the wing of the HALO aircraft. Right: MTP sensor unit in the lab. Marked with numbers are the radiometer unit (1), the hot calibration target (2), the rotating mirror (3) and the electronic unit (4), which contains various temperature sensors such as the scanning unit temperature or the pod air temperature sensor.
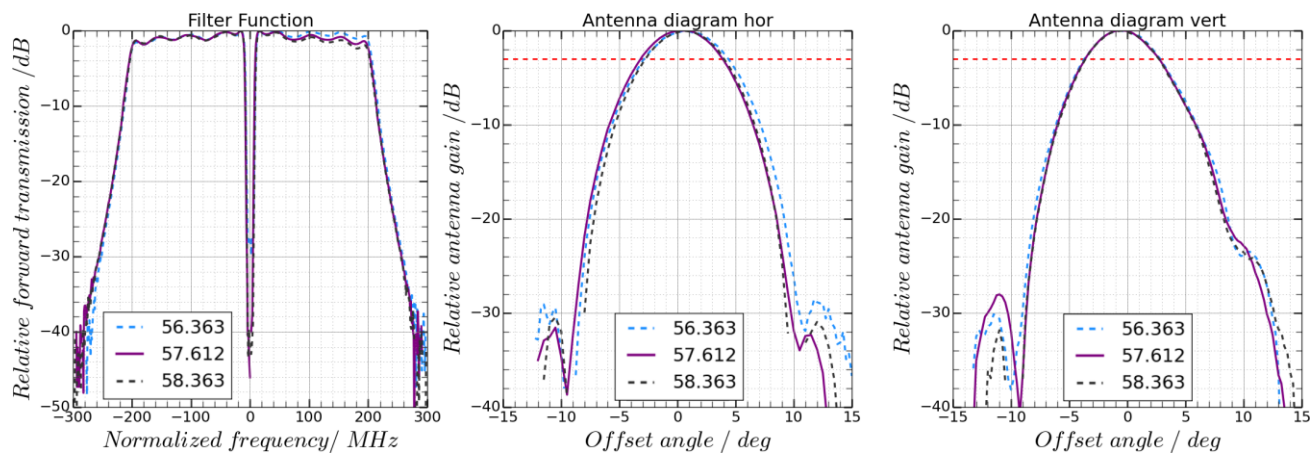
Figure 2: ~~DLR~~HALO-MTP Filter functions (left panel) and antenna diagram of the horizontal (middle) and vertical (right) plane recorded at standard measurement frequencies. Red dashed lines indicate the half maximum value. All data are normalised so that the maximum value shown is 0.

5
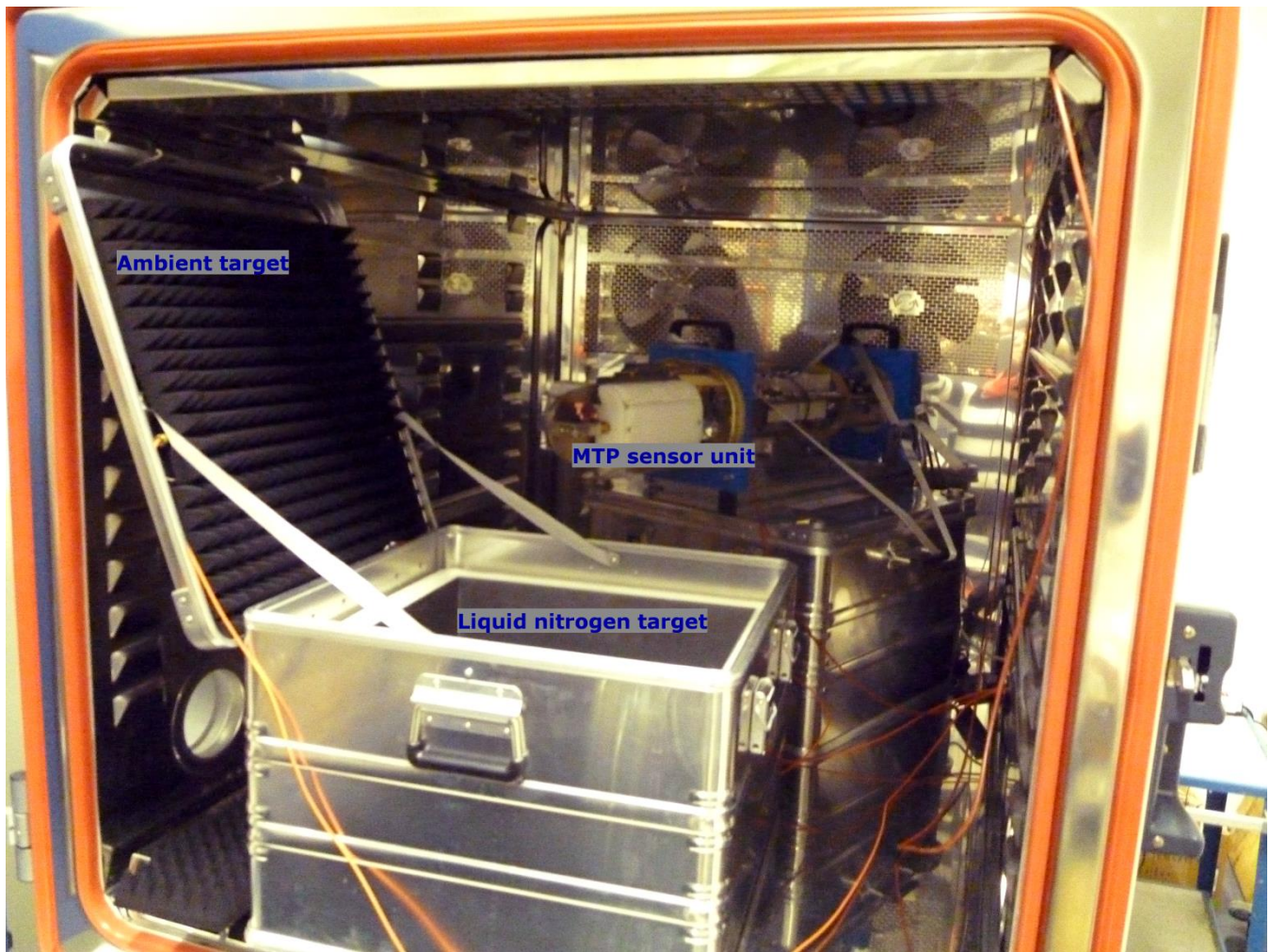
Figure 3: ~~DLR~~HALO-MTP inside the cold-chamber. For the measurements, the box containing the liquid nitrogen and the ambient target was rotated to face towards the MTP sensor unit. A second microwave absorber was placed on the ceiling of the chamber to function as second ambient target.
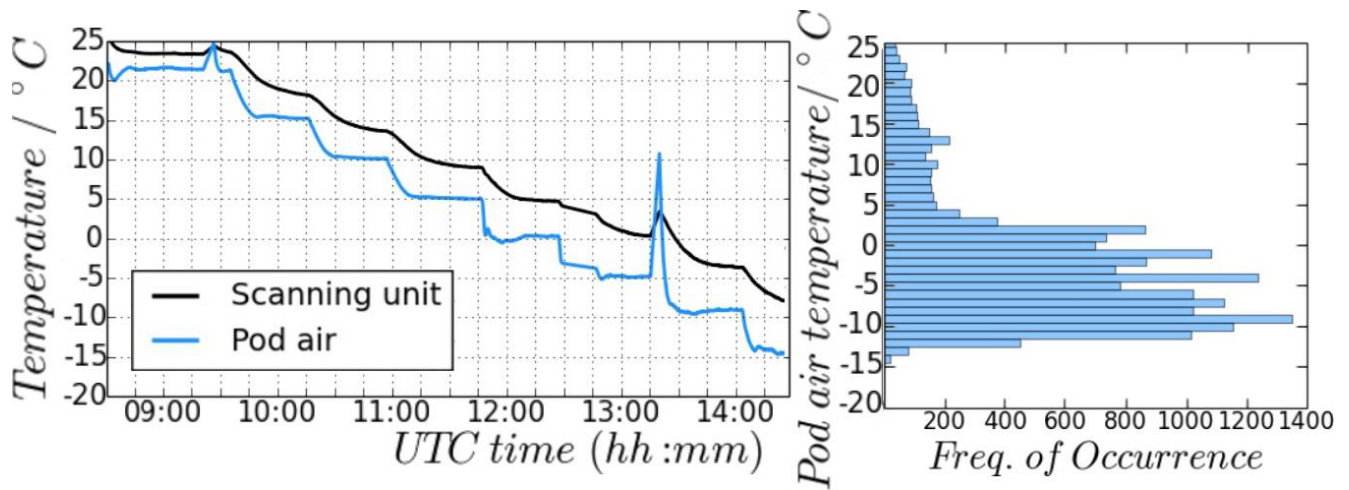
5

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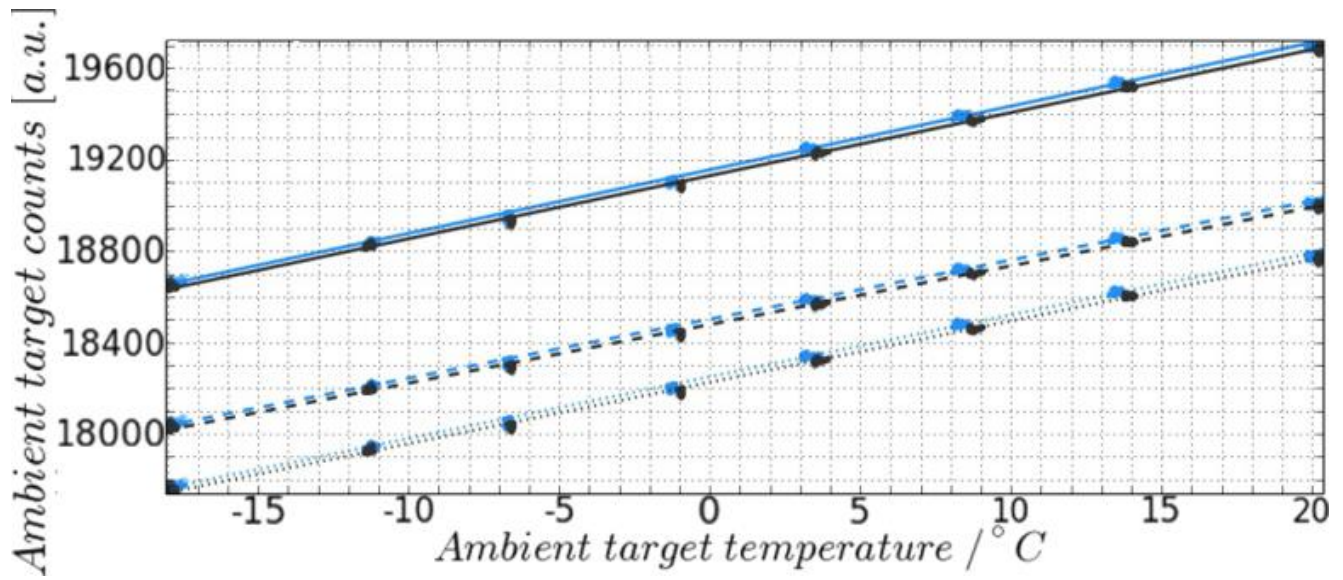
Figure 5: Ambient target temperature vs. measured signal (counts) of the two ambient targets for all three standard ~~LOs offrequency channelsof~~ the ~~DLR~~HALO-MTP (different line styles). Different line colours correspond to the measurement of the two individual ambient targets.

5
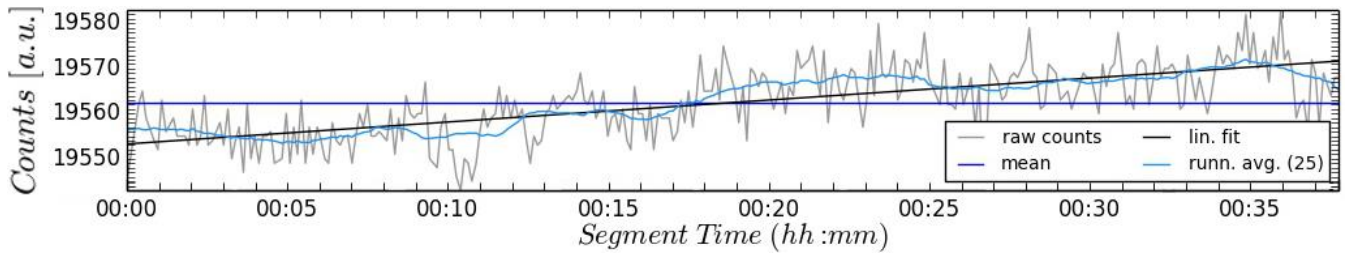
Figure 6: Measured signal (grey line) at ~~LO~~ 56.363 GHz while looking at the ambient target inside the cold chamber as well as a running average (light, blue line), mean value (blue line) and linear fit (black line). The corresponding brightness temperature change in the linear fit during the segment is about 0.8 K.
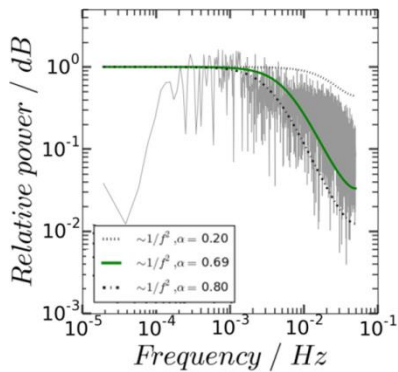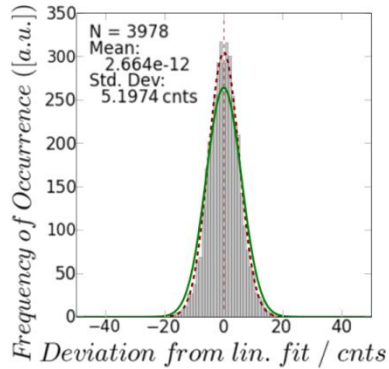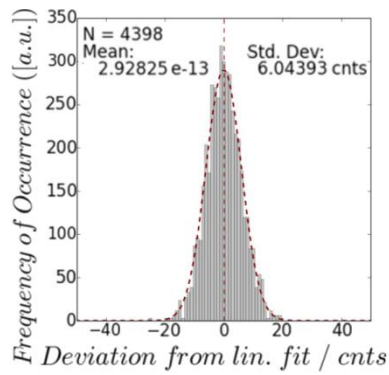
5

Figure 7: ~~DLR~~HALO-MTP noise figure at ~~LO~~ 56.363 GHz. Figures of the other two standard ~~LOs~~frequency channels look similar. Fit parameters for all three ~~LOs~~frequency channels are summarised in Table 2. Red dashed line: Gauss-fit to data. Top: Laboratory measurements in cold chamber. Middle: derived from ML CIRRUS flight data. Green line: Ideal Gauss function with the mean at 0.0 cnts and 6 cnts standard deviation~~,~~, as implied by the cold chamber noise figure. 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.
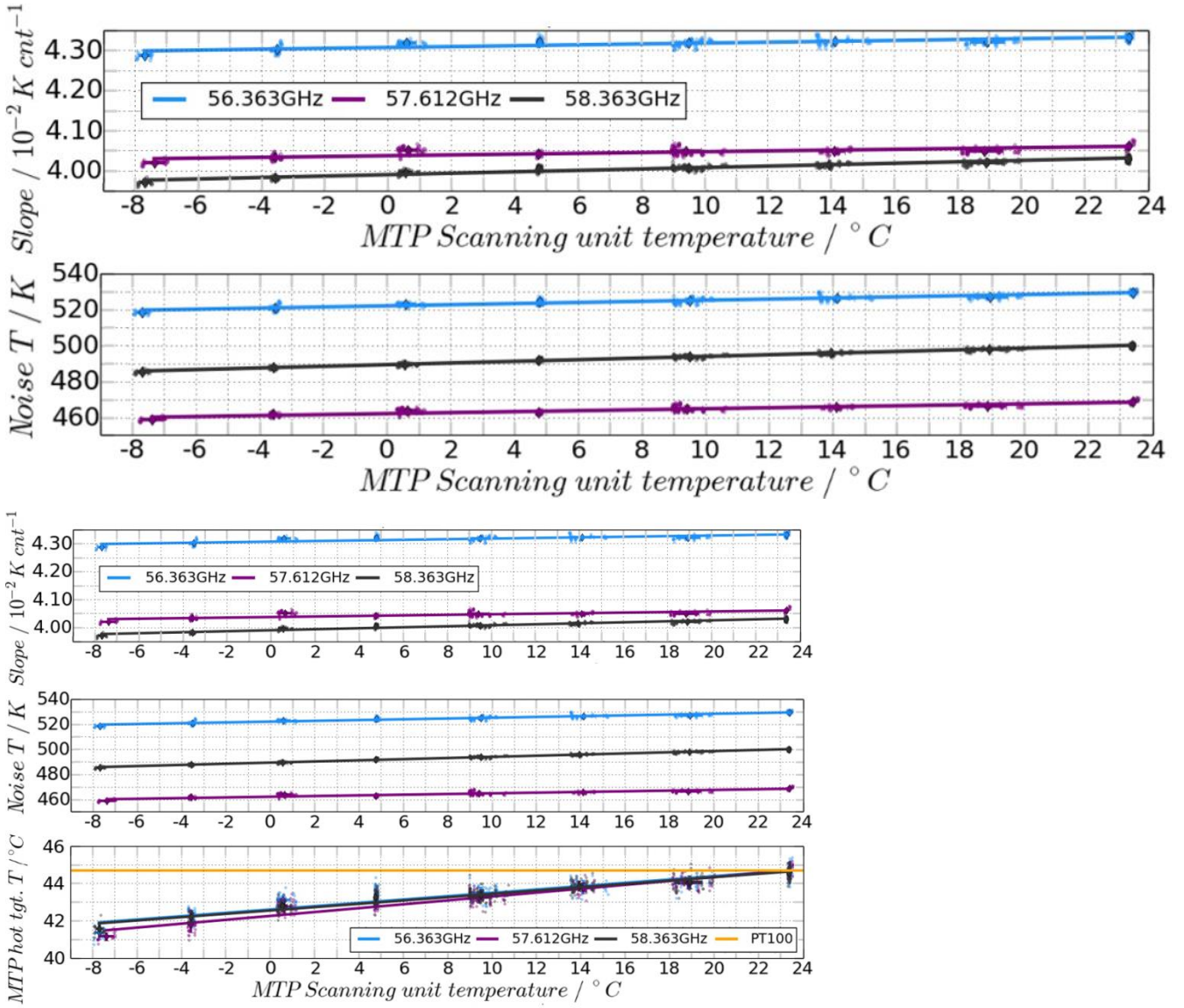
Figure 8: Calibration parameters calculated from hot-cold calibration for standard ~~LOs~~frequency channels during cold chamber measurements. Top: slope of calibration line. ~~Bottom~~Middle: Receiver noise temperature (T_R).
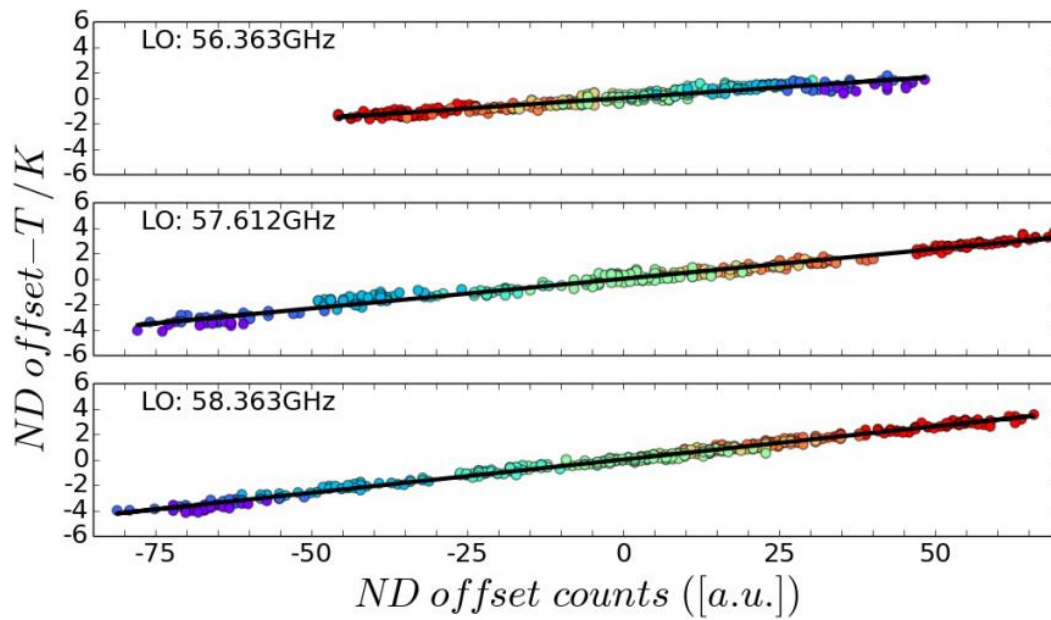
Figure 9: Calculated ND offset temperature for occurring ND count offsets (mean removed from both values). Colour coding: MTP scanning unit temperature (Blue: colder – red: warmer). Black line: Linear fit, linking ND offset temperature to offset counts.
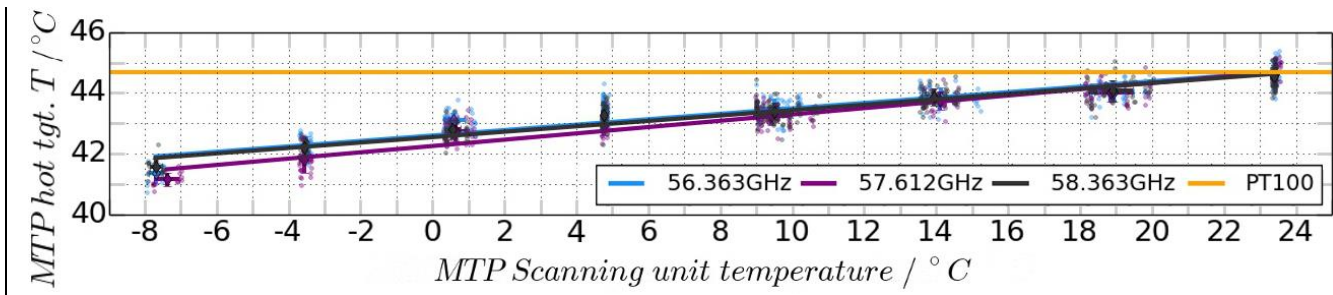
5

Figure 10 Bottom: Calculated hot target ~~BTs~~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.
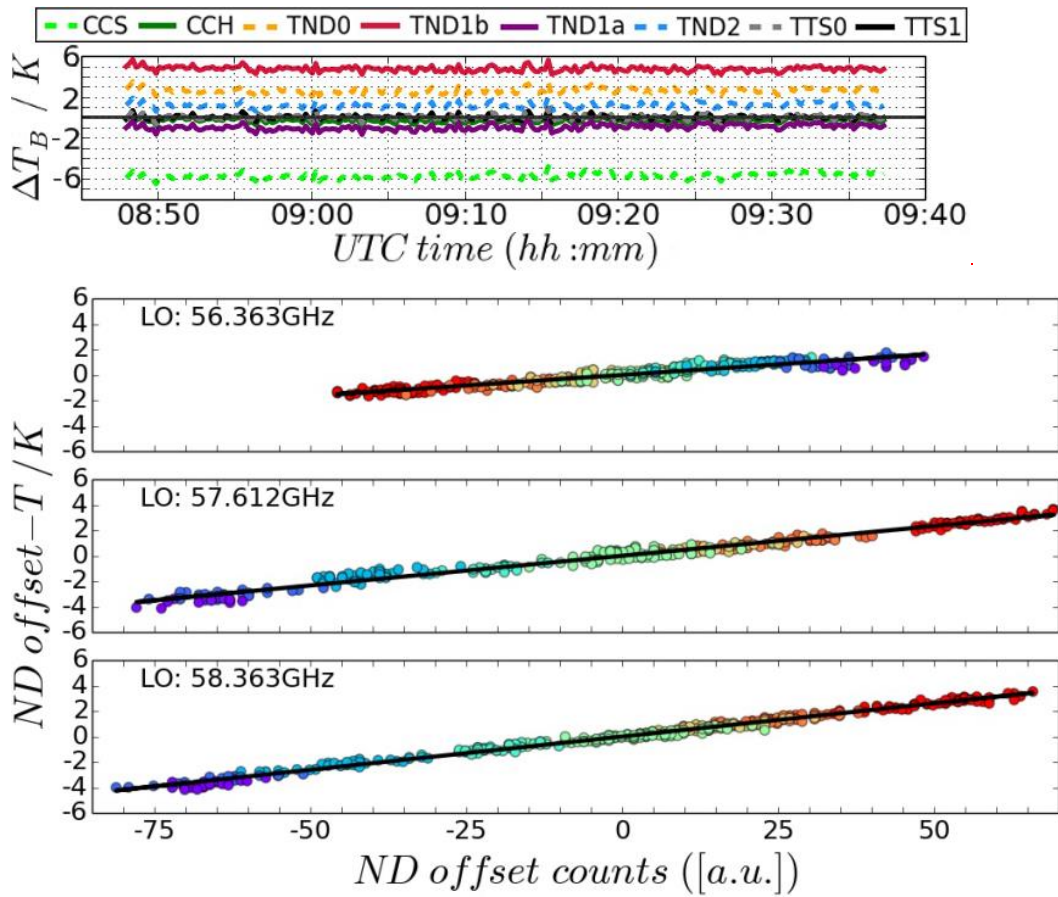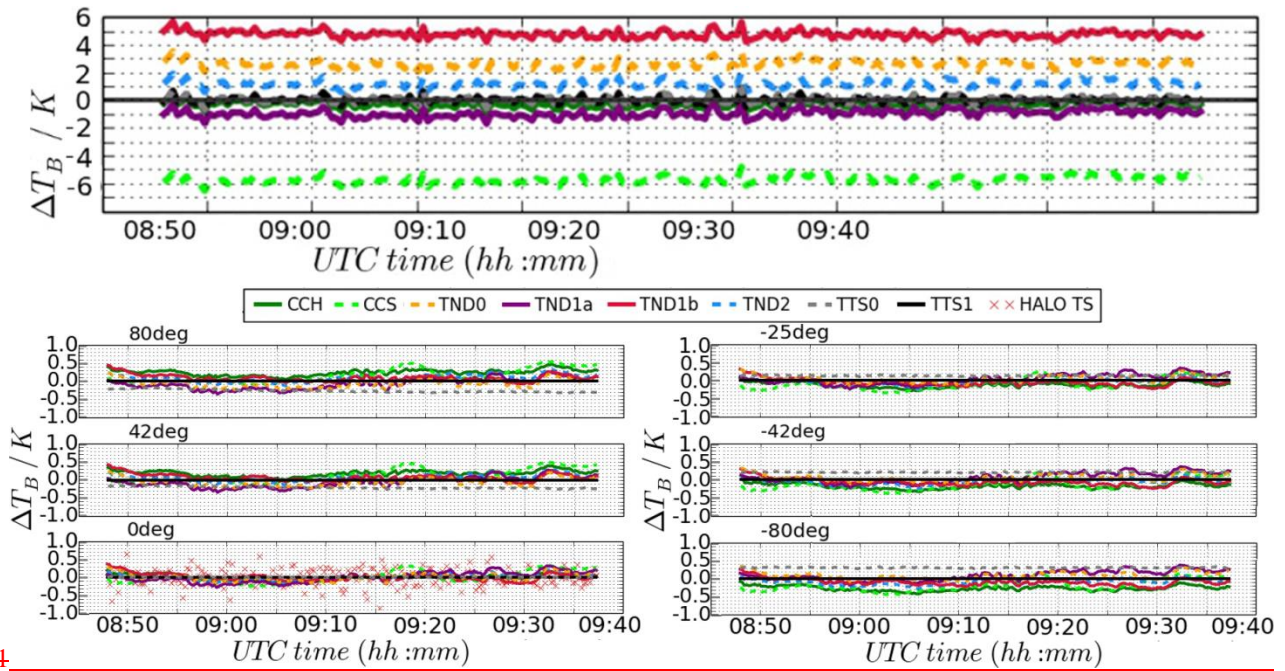
5

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.

Figure 10: Top panel: Difference between ~~BTs~~HALO static temperature and TBs derived with the eight calibration methods defined in Table 6 ~~and the BTs of 'TTS1'~~ derived from the horizontal measurements at 56.363 GHz during one segment of an ML CIRRUS mission flight. ~~.~~
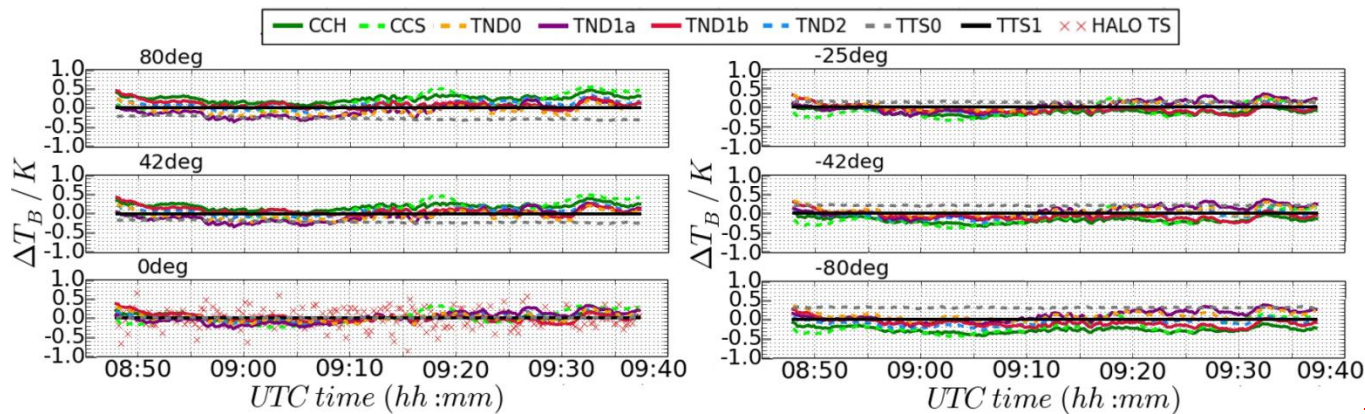
5

Figure 12: ~~Difference between offset-corrected BTs derived with the eight calibration methods defined in Table 6 and the offset-corrected BTs of 'TTS1'~~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.
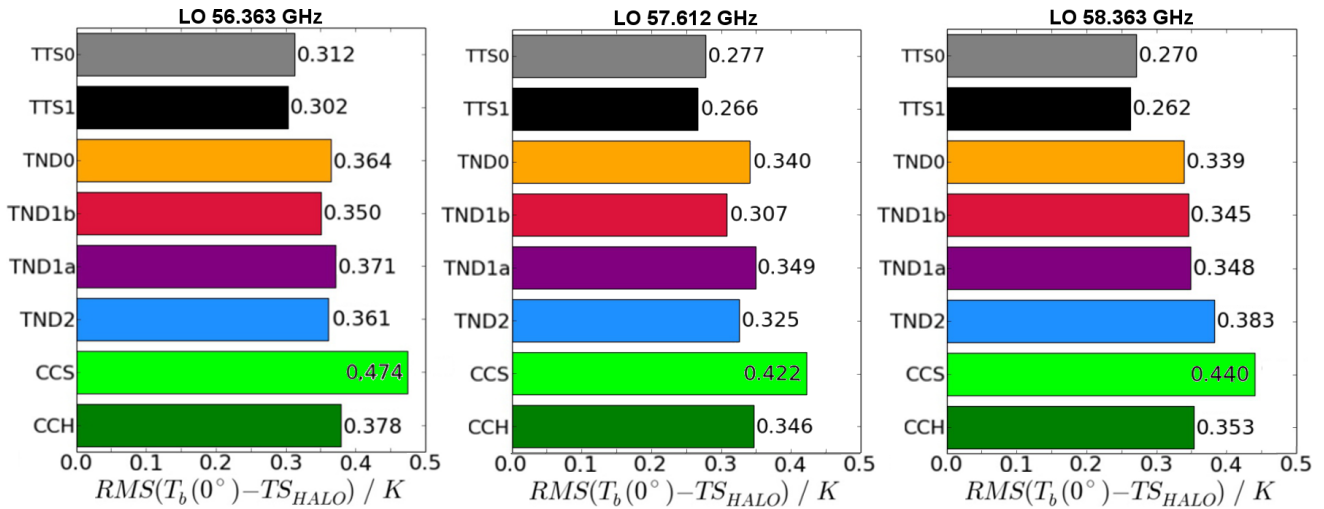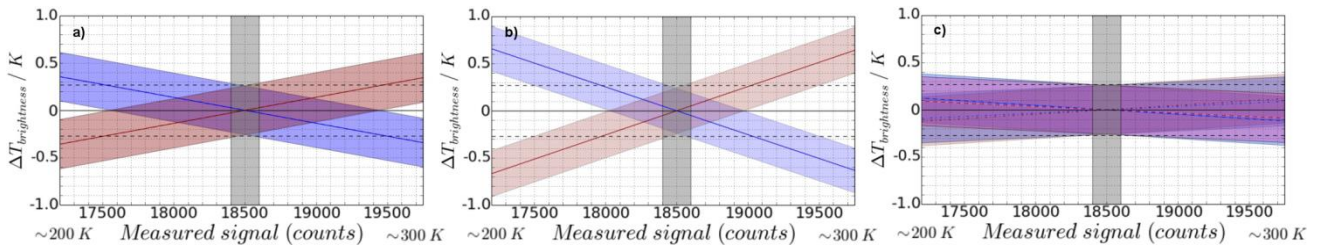
5

Figure ~~13~~11: RMS difference between HALO TS and ~~BTs~~TBs, derived from MTP measurement signal at limb-viewing angle 0∘ at the three standard ~~LO~~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 ~~14~~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.
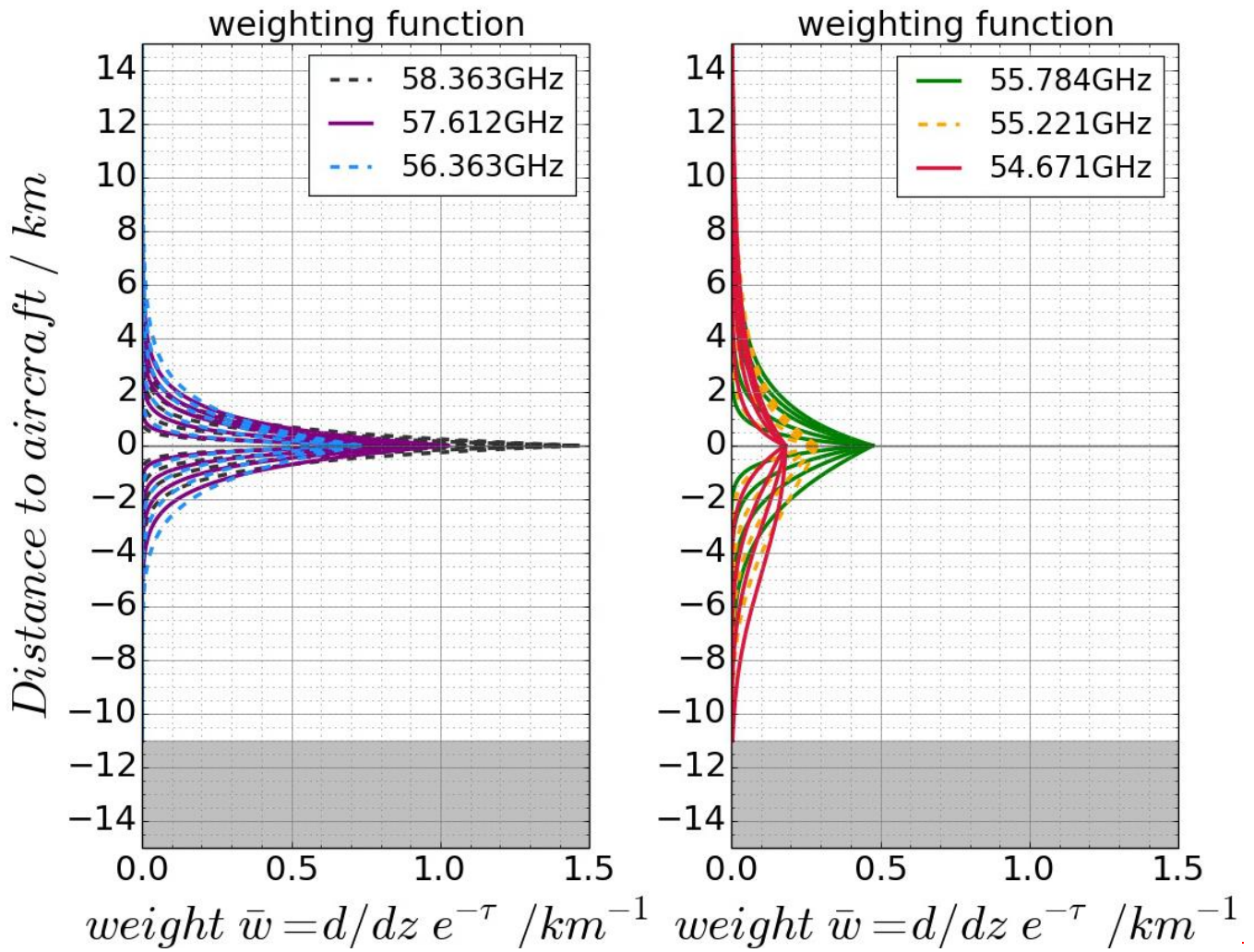
Figure 15: WFs of the MTP LOs (each individual line corresponding to a different viewing angle), calculated at aircraft altitude of 11 km. The $\bar{w}$ indicates averaging over all contributing frequencies within filter transmission range. Shown are the three standard LOs (left), and three possible LOs (right) to be used in a new measurement strategy of the DLR-MTP. Grey areas at the bottom: altitude range that would be below the surface.

Figure 16: Simulated BTs at LOs 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 LOs of the MTP; dashed vertical lines: strong lines that could be used as new MTP LO. Shading around vertical lines: MTP filter width.
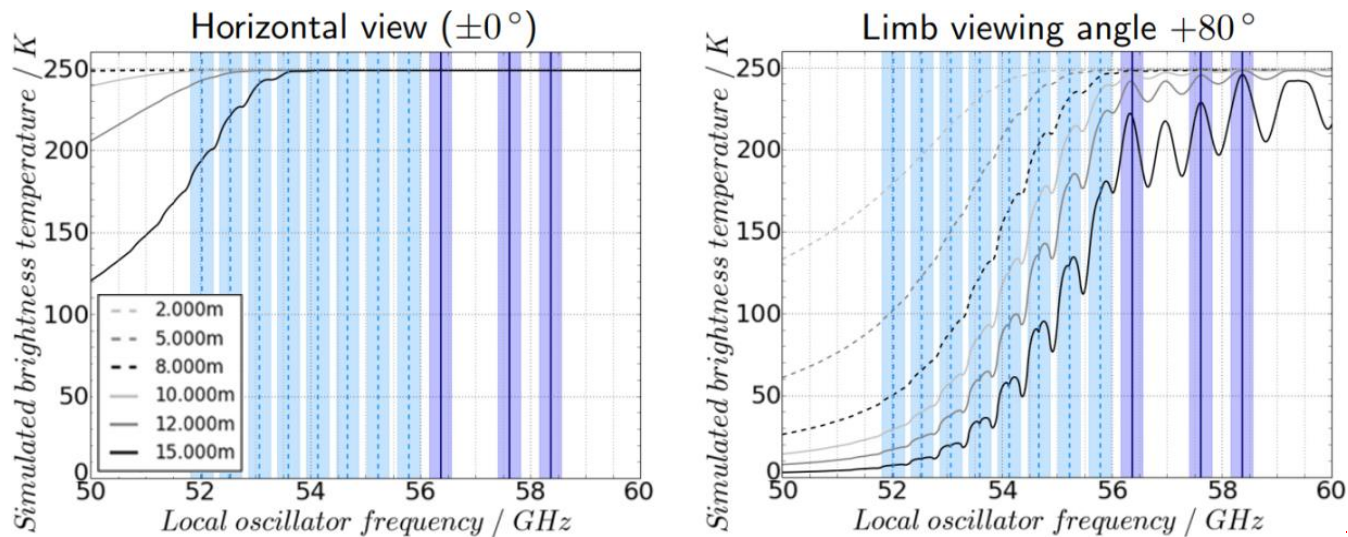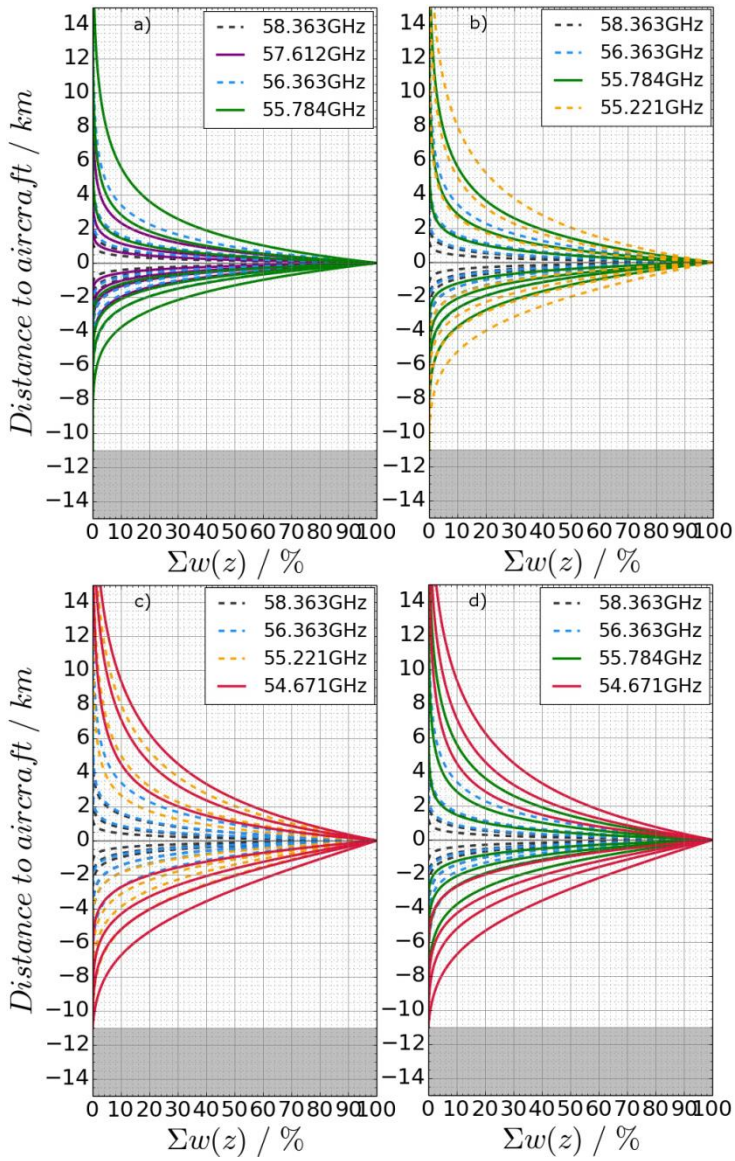
Figure 17: Cumulated weights of the MTP LOs (each individual line corresponding to a different viewing angle and frequency combination), calculated at aircraft altitude of 11 km. Shown are possible new measurement strategies, as mentioned in Table 9 a)'8E4LOa' b)'8E4LOc', c)'8E4LOb1', and d)'8E4LOb2'. Grey areas at the bottom: altitude range that would be below the surface.

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 to have a beam width of 7.5°; 360° rotatable. Steper 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 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.. ~~41~~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 **(N = 4398)**	Mean: 2.92825e-13	Mean: -1.5046e-12	Mean: -2.64535e-12
	Std. Dev.: 6.04393 cnts	Std. Dev.: 6.0963 cnts	Std. Dev.: 6.2264 cnts
Gauss fit: flight data **(N = 3978)**	Mean: 2.664e-12	Mean: 9.1452e-14	Mean: 3.7587e-13
	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: MTP instrument noise characteristics at each of the three standard ~~LO frequencies~~frequency channels.

LO	Scanning unit temperature					Hot target counts c_{hot}				
	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 radiative transfer (RT) simulations 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, $\mathcal{T}(v) = \exp(-\tau(v))$, characterizes the ratio of outgoing to incoming radiation traversing an atmospheric layer with path coordinate s. It is expressed through the optical depth $\tau(v)$ defined as the integral of the absorption coefficient (α), which depends on the frequency (v), the (path-dependent) atmospheric pressure ($p(s)$) and temperature ($T(s)$) of the layer within the plane-parallel atmosphere:

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

For brevity, in the following, the path-dependency of α is expressed as $\alpha(v, s)$. 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:

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

For RT calculations the Python scripts for Computational Atmospheric Spectroscopy (Py4CAtS[2]; 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 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 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

---

[2] available at http://atmos.eoc.dlr.de/tools/Py4CAtS/

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 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

82

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 $\pm80°$. 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 ($\pm80°$ relative to the horizon):

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

The relative path lengths to the $\pm80\circ$ 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 $l_{rel80°} \approx 1.5$. 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 $\pm14°$, $\pm30°$, $\pm41°$, and $\pm80°$, taking into account 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 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 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.
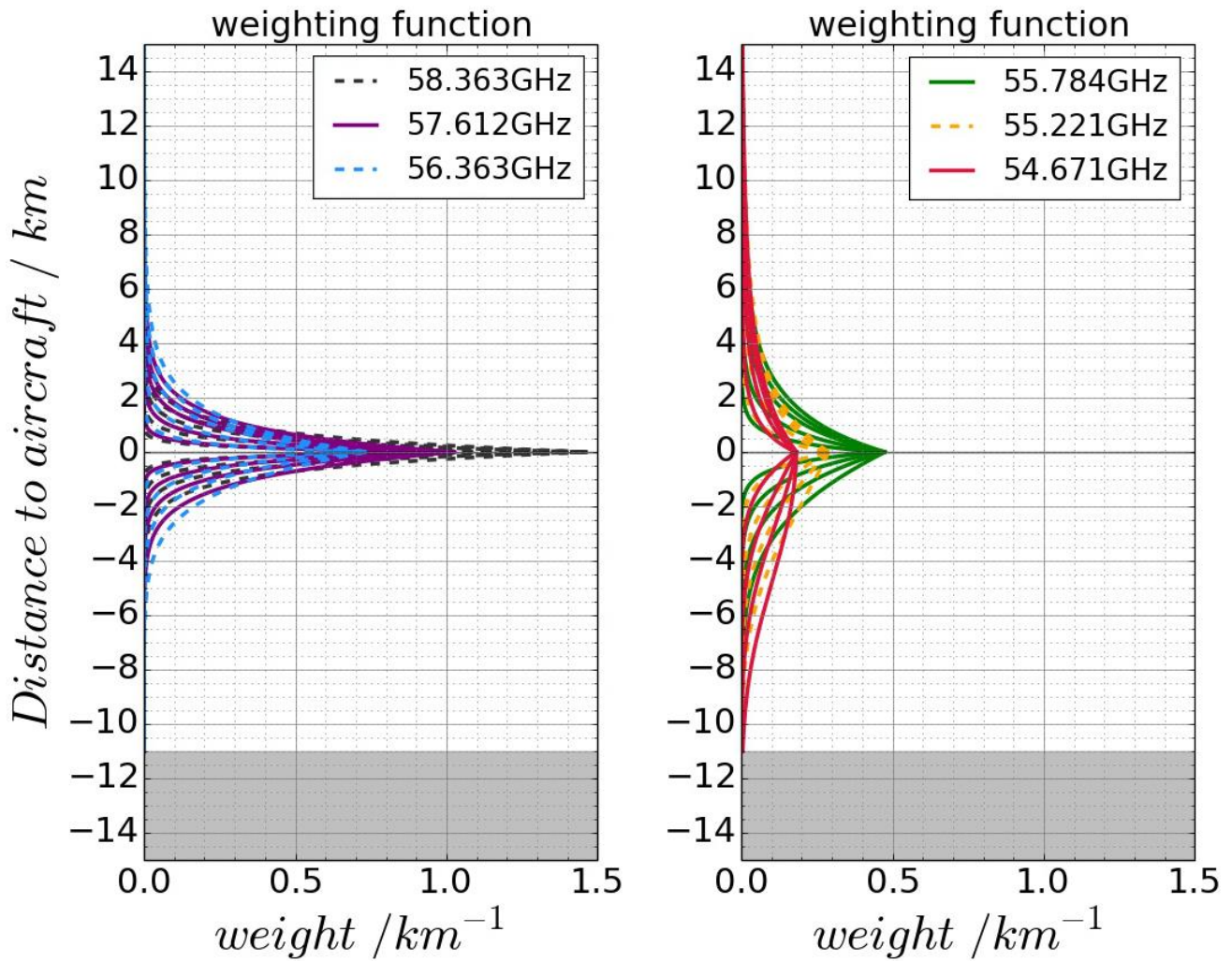
Figure A1: WFs (averaged over all contributing frequencies within filter transmission range) of the MTP frequency channels (each individual line corresponding to a different viewing angle), calculated at aircraft altitude of 11 km. Shown are the three standard frequency channels (left), and three possible 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.
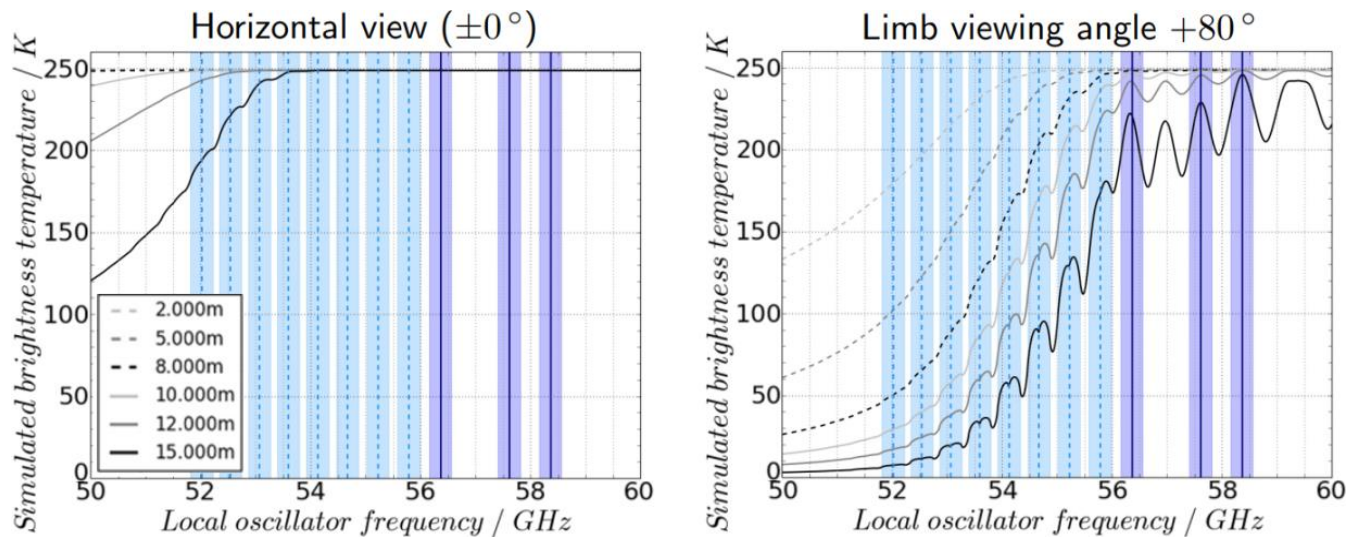
Figure A2: Simulated TBs at frequency channels between 50 GHz and 60 GHz at different flight altitudes (different line styles) and at horizontal viewing angle (left panel), and at near-zenith angle (right panel). Solid, vertical lines: Standard frequency channels of the MTP; dashed vertical lines: strong 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.
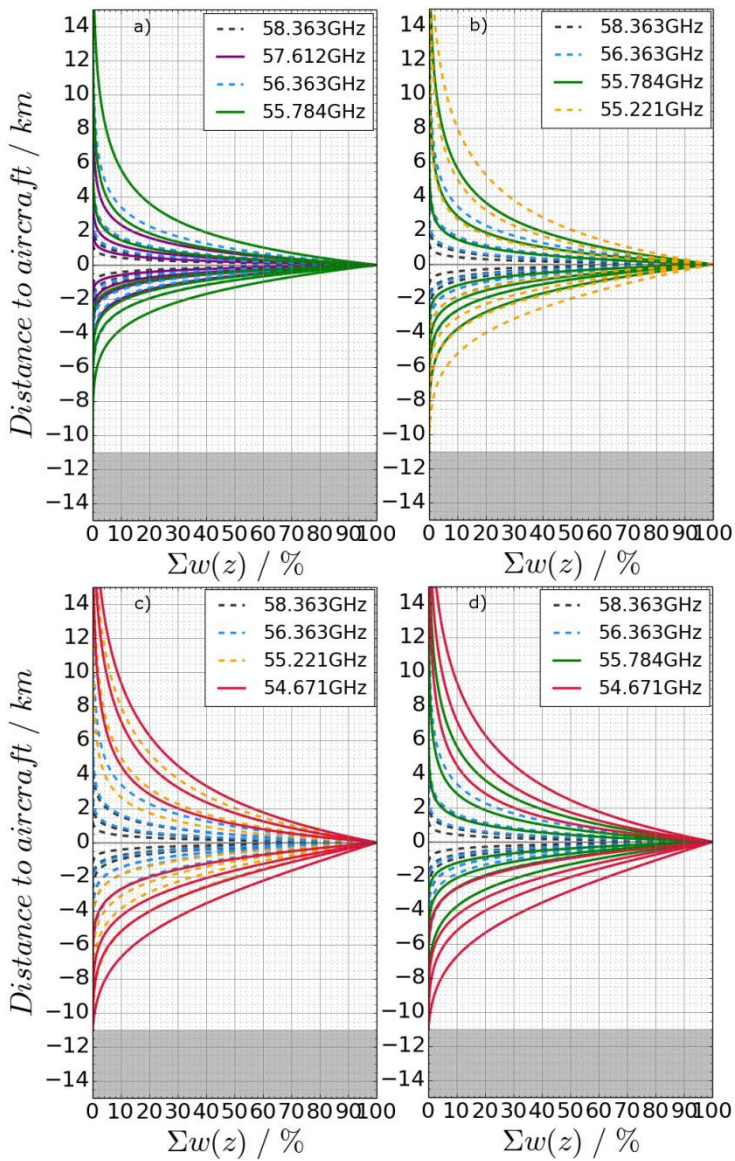
5

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 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 8A1: 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 9A2: Proposed measurement strategies for future missions of the ~~DLR~~HALO-MTP.