



# Investigation of ground-based microwave radiometer calibration techniques at 530 hPa

G. Maschwitz<sup>1</sup>, U. Löhnert<sup>1</sup>, S. Crewell<sup>1</sup>, T. Rose<sup>2</sup>, and D. D. Turner<sup>3</sup>

<sup>1</sup>University of Cologne, Institute for Geophysics and Meteorology, Cologne, Germany

<sup>2</sup>Radiometer Physics GmbH, Meckenheim, Germany

<sup>3</sup>National Severe Storms Laboratory, Forecast Research and Development Division, Oklahoma, USA

Correspondence to: G. Maschwitz (maschwitz@radiometer-physics.de)

Received: 29 October 2012 – Published in Atmos. Meas. Tech. Discuss.: 30 January 2013

Revised: 6 August 2013 – Accepted: 16 August 2013 – Published: 15 October 2013

**Abstract.** Ground-based microwave radiometers (MWR) are becoming more and more common for remotely sensing the atmospheric temperature and humidity profile as well as path-integrated cloud liquid water content. The calibration accuracy of the state-of-the-art MWR HATPRO-G2 (Humidity And Temperature Profiler – Generation 2) was investigated during the second phase of the Radiative Heating in Underexplored Bands Campaign (RHUBC-II) in northern Chile (5320 m above mean sea level, 530 hPa) conducted by the Atmospheric Radiation Measurement (ARM) program conducted between August and October 2009. This study assesses the quality of the two frequently used liquid nitrogen and tipping curve calibrations by performing a detailed error propagation study, which exploits the unique atmospheric conditions of RHUBC-II. Both methods are known to have open issues concerning systematic offsets and calibration repeatability. For the tipping curve calibration an uncertainty of  $\pm 0.1$  to  $\pm 0.2$  K (K-band) and  $\pm 0.6$  to  $\pm 0.7$  K (V-band) is found. The uncertainty in the tipping curve calibration is mainly due to atmospheric inhomogeneities and the assumed air mass correction for the Earth curvature. For the liquid nitrogen calibration the estimated uncertainty of  $\pm 0.3$  to  $\pm 1.6$  K is dominated by the uncertainty of the reflectivity of the liquid nitrogen target. A direct comparison between the two calibration techniques shows that for six of the nine channels that can be calibrated with both methods, they agree within the assessed uncertainties. For the other three channels the unexplained discrepancy is below 0.5 K. Systematic offsets, which may cause the disagreement of both methods within their estimated uncertainties, are discussed.

## 1 Introduction

Passive remote sensing instruments are widely used to retrieve atmospheric state variables. Measurements along the absorption lines of atmospheric gases can provide vertically resolved profiles of the gas concentration and/or temperature. Therefore, a realistic simulation of atmospheric absorption and emission is essential for the development of retrievals for atmospheric state variables. As radiative transfer models employ absorption line parameters that are obtained from laboratory measurements, field measurements are needed for independent evaluation. The aim of the Radiative Heating in Underexplored Bands Campaign (RHUBC-II) (Turner and Mlawer, 2010), which is part of the Atmospheric Radiation Measurement (ARM) program (Stokes and Schwartz, 1994), is to evaluate and improve existing absorption models. In addition to frequent radiosonde profiles of temperature, pressure and humidity, RHUBC-II data offer a unique set of accurate radiation measurements of the middle-to-upper troposphere. This paper concentrates on the microwave spectral region, which is widely used by space-borne and ground-based radiometers. For RHUBC-II this spectral range is covered by the passive microwave radiometer HATPRO-G2 (Humidity And Temperature Profiler – Generation 2). The radiometer has seven K-band channels along the rotational water vapor line at 22 GHz to retrieve information about the atmospheric water vapor and liquid water content. Another seven V-band channels along the oxygen absorption complex at 60 GHz are used to retrieve temperature profiles.

Ground-based temperature profiling using V-band frequencies was first suggested by Westwater (1965) and is now

performed operationally at several sites worldwide (Guldner and Spänkuch, 2001). However, compared to radiosonde profiles, temperature profiles derived from radiometric measurements can show biases up to 1 K (Liljegren, 2002; Löhnert and Maier, 2012). It is still unknown whether these originate from uncertainties in the oxygen absorption line parameters or from inaccuracies in the radiometer's absolute calibration. On the one hand, Cadeddu et al. (2007) evaluate oxygen absorption characteristics around 60 GHz by radiometer measurements and test different sets of absorption coefficients. They show that different sets of oxygen absorption line parameters lead to retrieved temperature profiles that may differ by more than 2 K. On the other hand, the accuracy of the commonly used liquid nitrogen (LN<sub>2</sub>) calibration is found to be 1–2 K (Liljegren, 2002). Similar issues are present in the K-band for the remote sensing of water vapor (Hewison, 2007). This makes clear that calibration and absorption model uncertainties are subject to the same level of uncertainty. Consequently, microwave radiometer calibration techniques need to be improved and well characterized before measurements may be used to evaluate current absorption models. Thus, the main goal of this paper is to evaluate two commonly used independent microwave radiometer calibration methods by carrying out a detailed error propagation by realistically considering multiple error sources.

The LN<sub>2</sub> calibration method is compared to the tipping curve method (Han and Westwater, 2000). Generally, all microwave channels can be calibrated using liquid nitrogen as a cold absolute standard. However, a successful tipping curve calibration requires a non-optically thick atmosphere at the frequency at stake. At sea level only K-band channels are transparent enough to be calibrated by this method. Fortunately, for the low pressure conditions at the RHUBC-II site, the tipping curve method can also be applied to the two most transparent V-band channels, allowing a first-time independent evaluation of the LN<sub>2</sub> calibration in the V-band.

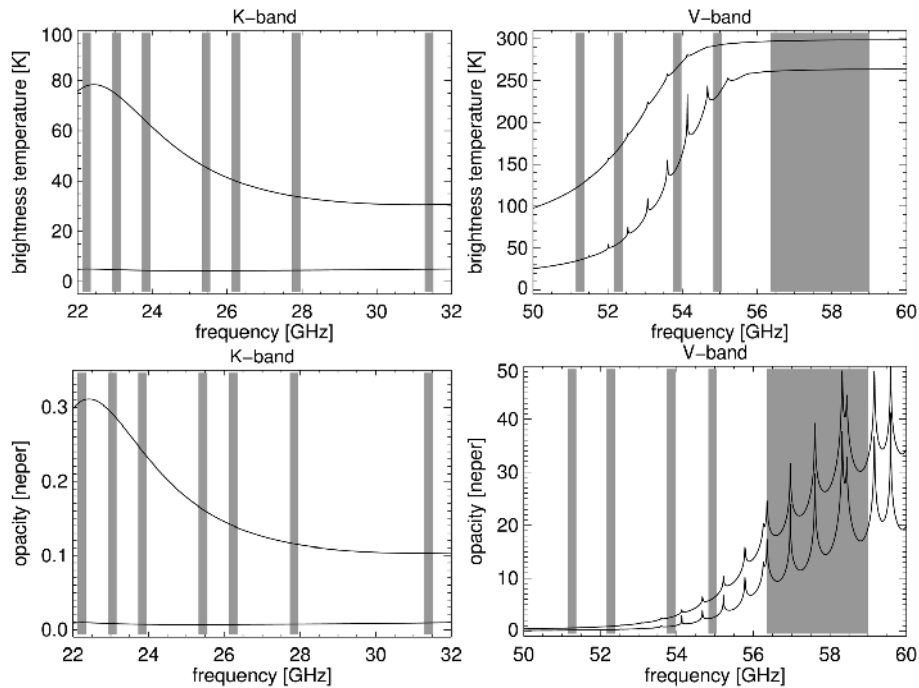
This paper is organized as follows: in Sect. 2 we briefly describe the RHUBC-II measurement campaign and the instrument that this study is based on. In the following Sect. 3 the theory behind calibration methods and the instrument used in this study are described, and Sect. 4 performs the actual calibration assessment. First, an error propagation study of the LN<sub>2</sub> calibration procedure is carried out in order to characterize the effects of different assumptions on the overall uncertainty. Specifically, these are the uncertainties attributed to the cold and hot target temperatures and emissivities, as well as to the detector non-linearity. Secondly, a similar procedure is carried out for the tipping curve calibration, where uncertainties such as pointing error, horizontal homogeneity and Earth curvature effects are discussed. Section 5 summarizes and compares the results of the two methods and Sect. 6 recommends steps to be taken in order to guarantee an optimal calibration result.

## 2 Measurement campaign

The primary focus of RHUBC-II was to characterize and improve the accuracy of gas absorption models using high spectral resolution radiance observations in spectral regions that are normally opaque at lower altitudes due to strong water vapor absorption and high air pressure (Turner and Mlawer, 2010). The measurement site was set up at 5320 m a.s.l. (above mean sea level) on the Cerro Toco next to the Chajnantor plateau (22.96° S, 67.77° W). The plateau is part of the Atacama Desert in northern Chile, which is one of the driest regions on Earth. Here, the amount of Precipitable Water Vapor (PWV) ranges from 0.1–5.0 mm (Rutllant, 1977).

RHUBC-II was conducted within the dry season between August and October 2009. The ARM program set up a Self-Contained Instrument Platform (SKIP) providing infrastructure and power supply for a variety of different instruments. Several radiometric instruments provide measurements almost across the entire infrared spectrum (Turner et al., 2012). The passive microwave radiometer HATPRO-G2 (Rose et al., 2005) took the data for this analysis and is described in detail below. The instrumentation was completed by an automatic weather station (AWS) and atmospheric profiles from 128 Vaisala RS92-K radiosondes launched during the campaign. During RHUBC-II PWV ranged from 0.2–1.5 mm with an average value of  $0.6 \pm 0.3$  mm. The synoptic station at the site recorded extraordinary constant surface conditions: a daily mean temperature of  $266.3 \pm 2.8$  K, an air pressure of  $532.2 \pm 2$  hPa, and a relative humidity of  $23 \pm 15$  %.

The HATPRO-G2 is a 14-channel total-power microwave radiometer measuring within the K- and the V-band, allowing high temporal resolution PWV, Liquid Water Path (LWP), humidity and temperature profile retrievals from brightness temperature measurements. It was manufactured by Radiometer Physics GmbH and is operated by the University of Cologne. In the K-band, atmospheric radiation is measured at seven channels along the wing of the water vapor absorption line at 22.35 GHz. Another seven V-band channels measure along the oxygen absorption complex centered around 60 GHz (Fig. 1). The receivers of each frequency band are designed as filter banks in order to acquire measurements at each frequency channel simultaneously. The channels have been designed with well characterized band-pass filters allowing precise frequency allocation of the signal. The eleven most transparent channels (K-band and V-band up to 54.94 GHz) have a half-power bandwidth of 230 MHz. For the three remaining channel frequencies the spectral gradient is low, so wider bandwidths of 600, 1000 and 2000 MHz provide higher precision at equal integration times. A parabolic mirror diameter of 250 mm gives an antenna half-power beam width for the channels along the water vapor line of 3.3–3.7° and 2.5–2.7° along the oxygen absorption complex. During RHUBC-II the radiometer was operated in a continuous scanning mode. It measured



**Fig. 1.** Highly spectrally resolved brightness temperature ( $T_b$ , top panels) and opacity ( $\tau$ , bottom panels), calculated by the Rosenkranz'98 absorption model (Rosenkranz, 1998), for the spectral range covered by HATPRO-G2. Spectral features are the water vapor absorption line at 22.24 GHz (left panels) and the oxygen absorption complex at 60 GHz (right panels). At 530 hPa, pressure broadening is reduced and single transitions can be recognized in the spectral wing of the oxygen absorption complex (lower lines). The upward-looking spectrum is plotted for a very dry (PWV = 0.3 mm) RHUBC-II radiosonde profile from 13 September 2009. For comparison the spectrum is calculated for sea-level conditions using a subtropical standard atmosphere (higher  $T_b$  and  $\tau$  values). HATPRO-G2 channels are illustrated by the width of band-pass filters (gray). Note that band-pass filters of the three most opaque channels overlap.

at symmetric elevation angles in a fixed azimuthal plane ( $70^\circ/250^\circ$ ) at elevations 90.0, 45.0, 30.0, 15.0, 9.6 and  $4.8^\circ$ . With an integration time of 1 s each scan takes about 30 s. In this study, these scans are used to apply the tipping curve calibration technique to the stored voltages of the nine low-opacity radiometer channels between 22.24 and 52.28 GHz.

### 3 Radiometer calibration

Calibrating the radiometer means to determine the relation between the detected voltage ( $U_{\text{det}}$ ) and the received power. The relation is derived by measuring against black body targets with a well-known physical temperature  $T_{\text{phys}}$ . The spectral radiance  $B_\nu$  that is emitted by the black body target is given by Planck's law:

$$B_\nu = \frac{2h\nu^3}{c^2} \frac{1}{\exp\left(\frac{h\nu}{k_b T_{\text{phys}}}\right) - 1} \quad (1)$$

in  $\text{W m}^{-2} \text{sr}^{-1} \text{Hz}^{-1}$ , with the Boltzmann constant  $k_b$  and Planck's constant  $h$ . Calibration procedures based on Eq. (1) provide measurements expressed in Planck equivalent brightness temperatures

$$T_b^{\text{PL}} = B_\nu^{-1}(I_\nu), \quad (2)$$

with the received spectral radiance  $I_\nu$ . In the following, given temperature values are always Planck equivalent. The relation between detected voltages and measured brightness temperatures  $T_b$  is expressed by the following set of calibration parameters:

$$U_{\text{det}} = g T_{\text{sys}}^\alpha, \quad (3)$$

with the detector gain  $g$ , the total system noise temperature  $T_{\text{sys}} = T_R + T_b$  (being the sum of the receiver noise temperature  $T_R$  and measured  $T_b$ ) and the non-linearity parameter  $\alpha$  (Radiometer Physics GmbH, 2011). At the beginning of every deployment period, the complete set of calibration parameters for HATPRO-G2 is determined by an  $\text{LN}_2$  calibration. Additionally, during operation noise diode calibrations (every 30 min) and hot load calibrations (every 5 min) are performed. For details on the procedures' input and output parameters confer Table 1. In the following the different procedures are introduced.

#### 3.1 Liquid nitrogen calibration

HATPRO-G2 uses an  $\text{LN}_2$  target that is mounted alongside the radiometer for calibration. The calibration target is

**Table 1.** Information on the standard calibration schedule during RHUBC-II. For both frequency bands the type, repetition period and number is given: liquid nitrogen calibration (LN<sub>2</sub>), tipping curve calibrations (TIP), noise diode calibrations (NOI) and hot load calibrations (HOT). Pointing: 90.0° = internal ambient target, 0.0° = LN<sub>2</sub> target; *i* = input parameter, *o* = output parameter.

K-band channels				
Type	LN <sub>2</sub>	TIP	NOI	HOT
Period	–	6 h	30 min	5 min
Number	1	97	4067	14 503
Pointing	–90°/0°	elev. scan	–90°	–90°
$U_{\text{det}}$	<i>i</i>	<i>i</i>	<i>i</i>	<i>i</i>
$T_{\text{H}}$	<i>i</i>	<i>i</i>	<i>i</i>	<i>i</i>
$T_{\text{C}}$	<i>i</i>	<i>i</i>	–	–
$\alpha$	<i>o</i>	<i>i</i>	<i>i</i>	<i>i</i>
$T_{\text{N}}$	<i>o</i>	<i>o</i>	<i>i</i>	–
$T_{\text{R}}$	<i>o</i>	<i>o</i>	<i>o</i>	<i>i</i>
<i>g</i>	<i>o</i>	<i>o</i>	<i>o</i>	<i>o</i>
V-band channels				
Type	LN <sub>2</sub>	TIP	NOI	HOT
Period	–	6 h	cont.	5 min
Number	1	–	–	14 503
Pointing	–90°/0°	–	scene	–90°
$U_{\text{det}}$	<i>i</i>	–	<i>i</i>	<i>i</i>
$T_{\text{H}}$	<i>i</i>	–	<i>i</i>	<i>i</i>
$T_{\text{C}}$	<i>i</i>	–	–	–
$\alpha$	<i>o</i>	–	<i>i</i>	<i>i</i>
$T_{\text{N}}$	<i>o</i>	–	<i>i</i>	–
$T_{\text{R}}$	<i>o</i>	–	–	<i>o</i>
<i>g</i>	<i>o</i>	–	<i>o</i>	<i>i</i>

observed from above using a reflector that is tilted by 45°. The load is filled with a microwave absorber to guarantee the target's black body properties. The target temperature  $T_{\text{C}}$  is the actual boiling temperature of LN<sub>2</sub>. The hot load is an internal black body at ambient temperature  $T_{\text{H}}$ , where venting ensures a homogeneous temperature distribution. In order to take the system non-linearity  $\alpha$  into account a 4-point calibration scheme is used (Fig. 2), which means that the radiometer measures at four calibration points to derive the calibration parameters  $g$ ,  $T_{\text{R}}$  and  $\alpha$  (Eq. 3) (Radiometer Physics GmbH, 2011). The determination of  $\alpha$  is realized by noise injection from an internal noise diode ( $T_{\text{N}}$ ). After integrating on the hot and cold target for 30 s each, the cycle is repeated with additionally injected noise.  $T_{\text{R}}$ ,  $g$ ,  $T_{\text{N}}$  and  $\alpha$  are then determined by solving the following system of equations:

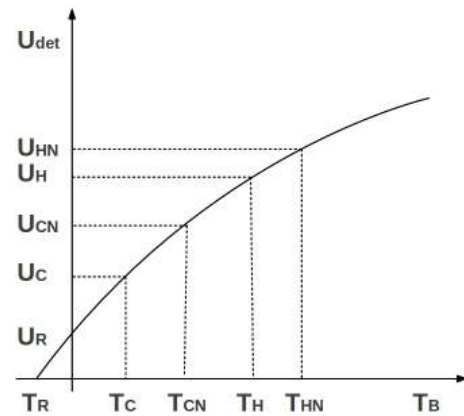
$$U_{\text{C}} = g (T_{\text{R}} + T_{\text{C}})^{\alpha}, \quad (4)$$

$$U_{\text{H}} = g (T_{\text{R}} + T_{\text{H}})^{\alpha}, \quad (5)$$

$$U_{\text{CN}} = g (T_{\text{R}} + T_{\text{C}} + T_{\text{N}})^{\alpha}, \quad (6)$$

and

$$U_{\text{HN}} = g (T_{\text{R}} + T_{\text{H}} + T_{\text{N}})^{\alpha}, \quad (7)$$



**Fig. 2.** Schematic illustration of a 4-point calibration. The detector voltage  $U_{\text{det}}$  is measured at four calibration points:  $U_{\text{C}}$ ,  $U_{\text{H}}$ ,  $U_{\text{CN}} = U_{\text{C}} + U_{\text{N}}$ ,  $U_{\text{HN}} = U_{\text{H}} + U_{\text{N}}$ .  $T_{\text{R}}$  is the system noise power,  $U_{\text{N}}$  results from injected noise by the internal noise diode ( $T_{\text{N}}$ ).  $T_{\text{C}}$  and  $T_{\text{H}}$  are known quantities,  $T_{\text{N}}$  and the non-linearity parameter  $\alpha$  are determined by the calibration.

with  $U_{\text{C}}$ ,  $U_{\text{H}}$ ,  $U_{\text{CN}}$ , and  $U_{\text{HN}}$  denoting radiometer voltages measured on the cold, ambient, cold-plus-noise and ambient-plus-noise references, respectively. Instead of using Planck spectral radiances  $B_{\nu}(T)$  within these calibration equations, Eq. (1) is expanded in terms of  $(h\nu/k_{\text{b}}T)$  and truncated after the first term. For HATPRO-G2's frequency range of 22 GHz <  $f$  < 58 GHz and temperature range of  $T_{\text{C}} < T_{\text{phys}} < T_{\text{H}}$  the truncation error of this approximation is negligible. Furthermore, Eqs. (4)–(7) are valid within HATPRO-G2's frequency range as long as the non-linearity  $\alpha$  is close to one. This condition is fulfilled by the HATPRO-G2 receivers. As a consequence, the target's physical temperatures can directly be inserted into Eqs. (4)–(7). The derived calibration parameters and the resulting brightness temperatures agree with the black body equivalent brightness temperatures  $T_{\text{b}}^{\text{PL}}$  (Eq. 2).

After a successful calibration,  $T_{\text{N}}$  from a well burned-in noise diode is stable enough to serve as a secondary calibration standard during operation. During RHUBC-II only one single LN<sub>2</sub> calibration was performed after radiometer setup on 11 August 2009. Therefore, in Sect. 4.1.5, we give a repeatability analysis of the calibration procedure using LN<sub>2</sub> calibration performed with the same radiometer at the Jülich Research Center.

### 3.2 Tipping curve calibration

An alternative to the LN<sub>2</sub> calibration is the so-called tipping curve calibration, which can be used to calibrate low-opacity radiometer channels. The general idea is to replace the LN<sub>2</sub> target by the cold clear sky. Han and Westwater (2000) discuss this method in detail and specify an absolute calibration accuracy of better than 0.5 K for K-band channels. In principle, the method uses opacity–air mass pairs from elevation

scans.  $T_b$  measurements are recalibrated in an iterative process, where the opacity  $\tau$  is calculated from the following equation:

$$\tau = \ln \left( \frac{B(T_{\text{mr}}) - B(T_{\text{back}})}{B(T_{\text{mr}}) - B(T_b)} \right), \quad (8)$$

with Planck spectral radiances  $B$  (Eq. 1).  $T_{\text{back}}$  ( $= 2.73$  K) denotes the cosmic background radiation and  $T_{\text{mr}}$  is the mean radiative temperature of the atmosphere, which is calculated independently (Sect. 4.2.1). The tipping curve calibration encompasses a large brightness temperature range and Eq. (8) is highly non-linear. Therefore, it is necessary to perform the tipping curve calibration in the power domain ( $B(T_b)$ ) and convert the power values to black body equivalent  $T_b$  by the inverse of Eq. (1) afterwards. Initial  $T_b$  values in Eq. (8) are set by using a prior calibration or an educated guess. Assuming clear sky conditions and a homogeneously stratified and non-opaque atmosphere, the opacity scales linearly with the air mass for low optical depths along the slant path. The slope of this linear relation is the zenith opacity  $\tau^{\text{zen}}$ . Zenith  $T_b^{\text{zen}}$  and  $\tau^{\text{zen}}$  are connected by

$$B(T_b^{\text{zen}}) = B(T_{\text{back}}) e^{-\tau^{\text{zen}}} + B(T_{\text{mr}}^{\text{zen}}) (1 - e^{-\tau^{\text{zen}}}). \quad (9)$$

$B(T_b^{\text{zen}})$  provides an updated cold reference at zenith. This procedure is repeated iteratively and converges to optimal calibration parameters  $g$  and  $T_R$  (Eq. 3), assuming that  $\alpha$  remains unchanged. For a perfect calibration the regression line would pass through the origin of the opacity–air mass diagram. However, deviations can result from horizontal atmospheric inhomogeneities within the scanned volume. Quality thresholds that guarantee the goodness of the fit are used to filter out tipping curve calibrations under inhomogeneous conditions.

Tipping curves are a standard calibration tool that is commonly used to calibrate K-band channels of ground-based radiometers. During RHUBC-II the manufacturer's internal procedure performed a tipping curve calibration of the K-band channels every 6 h. For this, the radiometer scanned towards an azimuth of  $50.0^\circ$  N at the following elevation angles:  $30.0$ ,  $33.3$ ,  $38.4$ ,  $45.6$ ,  $56.4$  and  $90^\circ$ . Throughout the campaign, 97 from 269 tipping curves passed the manufacturer's internal quality checks.

V-band channels are generally not included in tipping curve routines, because at sea level the opacity of these channels is too high. However, at 530 hPa, HATPRO-G2 channels at 51.26 and 52.28 GHz are transparent enough to apply the tipping curve calibration. Zenith  $T_b$  is below 40 and 60 K, respectively (Fig. 1) – well below the boiling temperature of  $\text{LN}_2$ . For the analysis in this study the tipping curve calibration is not only applied to the K-band channels but also the first two V-band channels (Sect. 4). Quality thresholds for the horizontal atmospheric homogeneity can be varied (cf. Sect. 4.2.5) and possible uncertainty sources can be investigated separately. The algorithm used within this analysis

has been verified with K-band results from the manufacturer's internal calibration procedure.

### 3.3 Noise diode calibration

$T_N$  is determined by every HATPRO-G2  $\text{LN}_2$  and tipping curve calibration (Sects. 3.1 and 3.2).  $T_N$  is stable enough to be used as a secondary calibration standard for several months (cf. Sect. 4.1.5). The stability of  $g$  for the V-band channels is guaranteed by injecting noise periodically – with a frequency of 10 Hz – while the radiometer is pointing to an arbitrary scene ( $T_{\text{scene}}$ ).  $T_{\text{scene}}$  can be extracted by using the fraction of the detected signal with and without noise.

Using  $T_R$  from a previous hot load calibration (cf. Sect. 3.4) gives a corrected detector gain  $g$ . Updating  $g$  with such high frequency is not necessary for the K-band channels, because the more stable gallium arsenide (GaAs) technology compared to indium phosphide (InP) technology (V-band) is used. Therefore, in the K-band a noise calibration is performed every 30 min to recalibrate  $T_R$ . The noise is injected while pointing to the ambient temperature target whereby  $T_R$  can be updated.

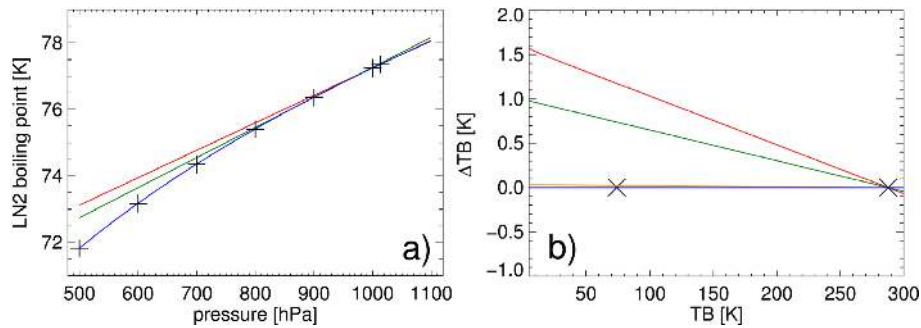
### 3.4 Hot load calibration

This calibration type uses the ambient hot load target as a reference. The hot load temperature  $T_H$  is known from a precision in-situ measurement within the target itself. HATPRO-G2 uses a second sensor to check for malfunction of the primary sensor. Generally, temperature gradients within the target may lead to a  $T_H$  that is not representative for the emitted radiance (McGrath and Hewison, 2001). For HATPRO-G2 such temperature gradients are sufficiently reduced ( $< 0.2$  K, Sect. 4.1.4), because the airflow from the internal blower covers the internal target. In order to correct for significant changes in HATPRO-G2's detector gain that occur on time scales longer than 5 min, the hot load target is reviewed every 5 min with an integration time of 4 s. For the V-band channels,  $g$  is known from continuous noise switching (Sect. 3.3) and thus  $T_N$  is used to update  $T_R$  during every hot load target calibration.

## 4 Calibration analysis

### 4.1 Liquid nitrogen calibration

Biases of 1–2 K are found when comparing simulated and measured  $T_b$  for radiometer channels along the oxygen absorption complex at 60 GHz (Liljgren, 2002; Löhnert and Maier, 2012). However, the question remains whether these biases originate from uncertainties in the  $\text{LN}_2$  calibration or in the oxygen absorption model. This bias would be too large for an absolute evaluation of absorption models because absorption model uncertainties are on the same order (Hewison and Cimini, 2006). Therefore, the absolute accuracy of



**Fig. 3.** Boiling point correction for LN<sub>2</sub>: different radiometer manufacturers have used different formulations for the pressure-dependent correction of the LN<sub>2</sub> boiling point. Absolute boiling points (a) and their impact on derived  $T_b$  (b,  $T_b$  reference from Eq. 15). Red: Eq. (12) (Radiometer Physics GmbH, 2011), green: Eq. (13) (Radiometrics, 2007), blue: new formulation (Eq. 15), orange: reference values (Span, 2000), crosses: calibration point  $T_C$  and  $T_H$ . (b) Impact on the derived  $T_b$  from differences in the assumed temperature of the cold calibration target ( $T_C$ ). Lines: LN<sub>2</sub> calibration simulations using the different corrections for the 23.04 GHz HATPRO-G2 channel, X: cold and hot calibration point.

HATPRO-G2 measurements is assessed by propagating uncertainties of the calibration parameters through the LN<sub>2</sub> calibration procedure. The contribution of each source of uncertainty is estimated by separately varying the individual input parameters for  $U_{\text{det}}$  (Eqs. 4–7). The total system noise temperature  $T_{\text{sys}}$ , when pointing to the LN<sub>2</sub> target, is modeled as

$$T_{\text{sys}} = (1 - r)T_{\text{tar}} + rT_{\text{cont}} + T_{\text{R}} + T_{\text{N}}, \quad (10)$$

with the signal emitted by the target itself ( $T_{\text{tar}}$ ), a contaminating component  $T_{\text{cont}}$ , which is reflected into the signal beam, the receiver noise  $T_{\text{R}}$ , and  $T_{\text{N}}$ , which is injected by an internal noise diode.  $r$  is the reflectivity of the calibration target:

$$r = \frac{(n - 1)^2}{(n + 1)^2}, \quad (11)$$

with  $n$  being the real part of the target's refractive index.  $r$  would be zero in case of a perfect black body target. For further information on reflectivities of radiometer calibration targets refer to Randa et al. (2005).  $T_{\text{N}}$  and  $T_{\text{R}}$  are determined by the calibration procedure (Eqs. 4–7). The uncertainty of  $T_{\text{N}}$  and  $T_{\text{R}}$  results from error sources associated with  $T_{\text{tar}}$ ,  $T_{\text{cont}}$  and  $r$ . The associated error sources are analyzed in the following.

An important error source of a calibration that uses an external LN<sub>2</sub> target is the formation liquid water on the radome due to condensation. This effect is not discussed here, because HATPRO-G2 uses a combined dew blower/heater system, which effectively prevents condensation during measurements and LN<sub>2</sub> calibrations.

#### 4.1.1 Boiling point correction

When calibrating against a liquid nitrogen target, the LN<sub>2</sub> boiling point serves as reference. At standard pressure

( $p_0 = 1013.25$  hPa) this is  $T_0 = 77.36$  K. However, the boiling point depends on the atmospheric pressure  $p$  and has to be corrected accordingly. For HATPRO-G2 the corrected boiling point is calculated from surface pressure supplied by the internal sensor. At 530 hPa the correction is several Kelvin. Any inaccuracy in the applied correction impacts  $T_b$  – especially for scenes close to or below the cold calibration temperature. However, different instrument manufacturers use different formulas for the pressure-dependent correction (Fig. 3):

$$T_C^{\text{RPG}} = T_0 - 0.00825(1000.00 - p), \quad (12)$$

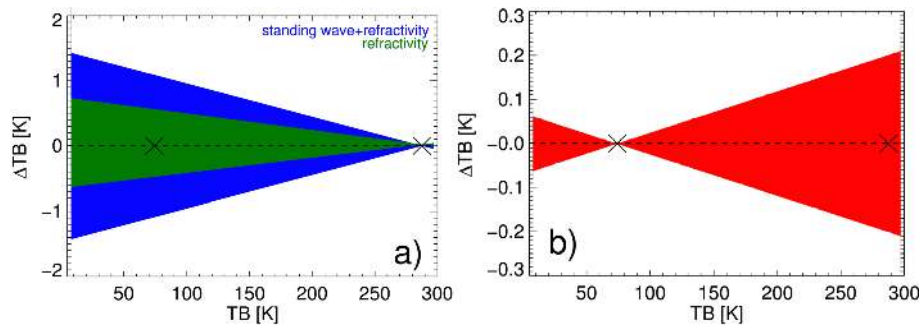
and

$$T_C^{\text{RAD}} = 68.23 + 0.009037 p, \quad (13)$$

with  $p$  in hPa.  $T_C^{\text{RPG}}$  and  $T_C^{\text{RAD}}$  refer to Radiometer Physics GmbH (2011) and Radiometrics (2007), respectively. The different formulations might be explained by the fact that the pressure dependency is often calculated for high technical pressures in the laboratory. At standard pressure the discrepancy between these formulations is still mostly negligible, however with decreasing pressure it increases and reaches 1.1 K at 530 hPa. In Fig. 3 the impact of the different boiling point corrections on the calibration is shown. The effect itself is not frequency-dependent, but the impact on measured  $T_b$  increases with decreasing channel opacity. It can be seen that moving away from the hot target increases the cold target's influence. At the LN<sub>2</sub> boiling point the discrepancy between the different corrections already exceeds 1 K. The proper pressure-dependent correction for the LN<sub>2</sub> boiling point can be calculated from the Clausius–Clapeyron equation:

$$\ln\left(\frac{p}{p_0}\right) = \frac{\Delta H_{\text{vap}}}{R} \left(\frac{1}{T_0} - \frac{1}{T}\right), \quad (14)$$

with the boiling point of liquid nitrogen  $T$  and the ideal gas constant  $R$ . Apart from the pressure Eq. (14) only depends



**Fig. 4.** (a) Uncertainty in derived  $T_b$  from uncertainties at the cold calibration point, blue: area of uncertainty resulting from the uncertainty for the refractive index of the  $\text{LN}_2$  surface ( $n = 1.20 \pm 0.03$  (Benson et al., 1983)), green: area of uncertainty for a single calibration due to resonances between the  $\text{LN}_2$  surface and the receiver. (b) Uncertainty in derived  $T_b$  from uncertainties at the hot calibration point, red: area of uncertainty from errors in the in-situ temperature measurement of the internal ambient target ( $\Delta T_H = \pm 0.2$  K), X: cold and hot calibration point. Note that the y axes of the two subplots have different limits.

on the heat of vaporization  $\Delta H_{\text{vap}}$ , which is constant for an ideal gas. Based on Eq. (14) a consistent formulation is now used:

$$T_C = \frac{710.5241}{9.185 - \ln\left(\frac{p}{1013.25}\right)}. \quad (15)$$

Using the atmospheric pressure during the  $\text{LN}_2$  calibration at RHUBC-II of  $p = 534.7$  hPa, Eq. (15) gives a boiling point of  $T_C = 72.2$  K. For RHUBC measurements this adapted boiling point decreases calibrated  $T_b$  values by up to 1.4 K compared to the original  $\text{LN}_2$  calibration (Eq. 12). Note that the National Institute of Standards and Technology (NIST) gives a boiling point of  $T_C = 72.3$  K at 534.7 hPa (Span, 2000).

#### 4.1.2 Refractive index of liquid nitrogen

The refractive index of the  $\text{LN}_2$  target,  $n_{\text{LN}_2}$ , determines the reflectivity  $r_{\text{LN}_2}$  of its surface (Eq. 11).  $n_{\text{LN}_2}$  is derived from laboratory measurements. Reesor et al. (1975) determine  $n_{\text{LN}_2} = 1.2$  for a frequency range of 18 to 26 GHz. The results are in agreement with several other experiments at frequencies between 0.5 GHz (Hosking, 1993) and 130 GHz (Vinogradov et al., 1967). Therefore,  $n_{\text{LN}_2} = 1.2$ , which corresponds to a reflectivity of  $r_{\text{LN}_2} = 0.82\%$  (Table 2), is assumed for all HATPRO-G2 channels. The contaminating signal  $T_{\text{cont}}$  (Eq. 10) is estimated to originate from the receiver ( $T_{\text{rec}} = 305$  K) and being reflected back at the  $\text{LN}_2$  surface. As  $T_{\text{rec}}$  is much higher than the  $\text{LN}_2$  target temperature even a small reflectivity results in a reflective component that has to be added to the cold target temperature. In this case, a reflective component is calculated to  $T_{\text{refl}} = 1.9 \pm 0.6$  K when an  $n_{\text{LN}_2}$  uncertainty of  $\pm 0.03$  is assumed (Benson et al., 1983). The resulting  $T_b$  uncertainty reaches 1.4 K for  $T_b = 5$  K, which is the minimum measured  $T_b$  within the K-band during RHUBC-II (Fig. 4). The effect decreases linearly with higher  $T_b$  values. For the most opaque channels in the V-band the uncertainty reduces to 0.1–0.2 K. Finally, it disappears at the hot calibration point.

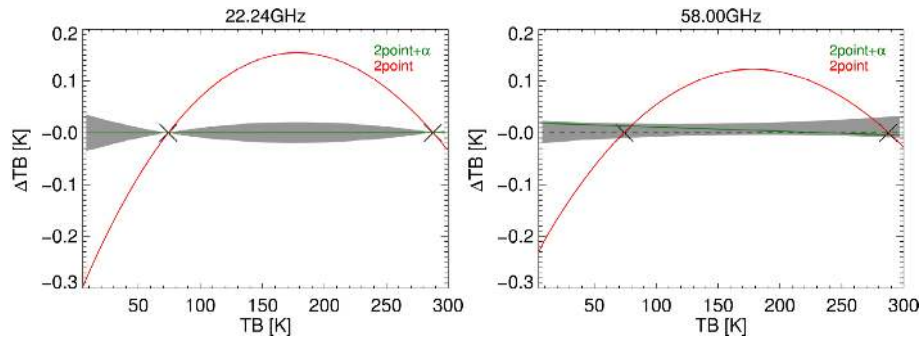
**Table 2.** Refractive indices ( $n_{\text{LN}_2}$ ) and reflectivities ( $r_{\text{LN}_2}$ ) of the liquid nitrogen target and the resulting reflective components ( $T_{\text{refl}}$ ).

$\nu$ GHz	$n_{\text{LN}_2}$	$r_{\text{LN}_2}$ [%]	$T_{\text{refl}}$ [K]
orig.	1.20	0.82	1.9
min	1.17	0.61	1.4
max	1.23	1.06	2.3

Additionally, the  $\text{LN}_2$  surface is exposed to the environment and is thus deformed by capillary waves and boiling bubbles. On the one hand, this can lead to frequency-dependent interferences resulting in variations in the reflectivity  $r_{\text{LN}_2}$  of the target (Shitov et al., 2011), because the deformation length scale is on the range of the received signal wavelength. These variations are on time scales well below the integration time on the  $\text{LN}_2$  target of 30 s. Therefore, they are assumed to compensate each other over the integration period. On the other hand, air bubbles on the  $\text{LN}_2$  surface reduce the density of the interacting surface. This effect will not be compensated by averaging and may improve the cold target's black body characteristics because  $r_{\text{LN}_2}$  is reduced. However, it is difficult to quantify this effect. Assuming that the density of the interacting  $\text{LN}_2$ –air volume is reduced by 1 % due to air bubbles would lead to a  $T_C$  reduced by 0.2 K.

#### Resonance effect

Continuous observations of the  $\text{LN}_2$  evaporating from the cold load were carried out at the Jülich Research Center. The observations were conducted with the same radiometer that was previously deployed at RHUBC-II. While the level of the  $\text{LN}_2$  declines, its distance to the receiver  $s$  increases and the resonance condition changes. The effect on  $T_{\text{tar}}$  (Eq. 10) can be interpreted as a small perturbation  $\text{res}(s)$ . Averaged over time,  $\text{res}(s)$  is zero ( $\langle \text{res}(s(t)) \rangle = 0$ ). Via  $r_{\text{res}}$  these standing waves contribute to the reflective component:



**Fig. 5.** The effect of the non-linearity parameter  $\alpha$  of the radiometer system and its uncertainty on calibrated  $T_b$  for two exemplary HATPRO-G2 channels. The exact 4-point scheme solution serves as the reference (zero line). Red: 2-point calibration ( $\alpha = 1$ , no noise injection), green: 2-point calibration using  $\alpha$  from the original 4-point calibration performed at the beginning of RHUBC-II, shaded area: uncertainty resulting from random uncertainty of  $\alpha$ .

$$T_{\text{refl}} = (r_{\text{LN}_2} + r_{\text{res}}(t)) (T_{\text{cont}} - T_C). \quad (16)$$

Note that  $r_{\text{res}}$  is unknown, because the reflectivity on the receiver end is not known. The corresponding amplitudes and phases of the observed signal are frequency-dependent and affect the uncertainty of the calibration point. The maximum uncertainty is estimated to be twice the amplitude of the oscillation observed at each channel, because the integration time within the LN<sub>2</sub> calibration is small compared to the oscillation periods. K-band channels show oscillation amplitudes of 0.1 to 0.6 K. In the V-band the noise level is generally higher, making it more difficult to detect oscillations, and may lead to a reduction of amplitudes, which are 0.1–0.3 K. For both receiver bands, amplitudes in the band's center show higher amplitude. Most probably this is caused by the isolator, which is located directly after the feed horn. The isolator is designed to cover the large bandwidth of the whole receiver. Its performance might slightly decrease towards the band edges. Figure 4 shows the impact of the resonances on calibrated  $T_b$  for 23.04 GHz. This channel has the largest uncertainty: for RHUBC-II measurements with  $T_b$  below 10 K the maximum uncertainty range reaches 1.2 K. The impact of the resonances can be diminished by integrating on the cold target for exactly one oscillation period. However, the periods are channel-dependent. Therefore, a more practical solution is to determine the calibration parameters from repeated calibrations while the LN<sub>2</sub> evaporates. The remaining uncertainty is due to the fact that only a limited number of repetitions is possible before the LN<sub>2</sub> has evaporated. In this study, the maximum uncertainty from resonances is assumed, because for RHUBC-II only a single cold load integration was performed.

#### 4.1.3 Non-linearity parameter

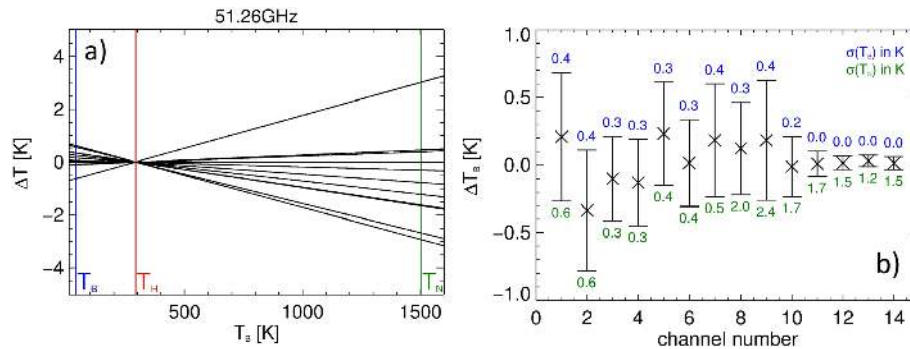
Concerning the non-linearity parameter  $\alpha$  there are two aspects to be analyzed: (1) the effect of  $\alpha$  on the calibration

characteristic ( $T_b = f(U_{\text{det}})$ ) and (2) its variability. The first point is addressed by simulating the LN<sub>2</sub> calibration with a classical 2-point calibration scheme. This scheme is a simplification of the 4-point scheme: it assumes a linear relation between  $U_{\text{det}}$  and measured  $T_b$  ( $\alpha = 1$ ) and no additional noise is injected. The impact on the calibration characteristic is given in Fig. 5. Apart from numerical effects, the impact diminishes at the two calibration points  $T_C$  and  $T_H$ . The impact on the calibration is slightly higher for V-band channels because  $\alpha$  is generally smaller than for the K-band channels. Nevertheless, the effect depends on  $T_b$ : the maximum difference of 0.3 K is reached at the lower end of the measured  $T_b$  range. Therefore, K-band channels are most affected. In any case, considering the detector non-linearity noticeably improves the calibration. The variability in  $\alpha$  is investigated by 27 LN<sub>2</sub> calibrations that were performed with HATPRO-G2 at the Jülich Research Center, Germany, between July 2010 and November 2011.  $\alpha$  cannot be determined independently, because it characterizes the non-linearity of the whole radiometer system. Within the considered time period, no significant drift of  $\alpha$  is found. Therefore, the standard deviation of  $\alpha$  over all 27 calibrations reflects the random uncertainty of the  $\alpha$  determination, which is 0.1–0.2% of the mean  $\alpha$  value of each channel. It can be concluded that  $\alpha$  is solely a frequency-dependent instrument property. For the K-band channels the variability is largest at 22.24 GHz and decreases with frequency. Figure 5 shows the impact on calibrated  $T_b$  when the original  $\alpha$  is varied within the range of uncertainty. Again, the effect is largest for small measured  $T_b$ . Combined with the higher variability of  $\alpha$ , K-band channels are most affected. Anyway, the effect does not exceed  $\pm 0.04$  K (Table 5) and is therefore negligible.

#### 4.1.4 Hot load temperature

All calibration types (Table 1) use the internal ambient target as a hot reference. The main source of uncertainty is the in-situ temperature measurement of the target. It is considered





**Fig. 6.** (a) Impact of variations in noise diode temperature  $T_N$  on  $T_b$  measurements for the 51.26 GHz channel. Vertical lines mark the mean brightness temperature  $T_b = 38.4$  K, the ambient target temperature  $T_H = 293.1$  K, and the average noise diode temperature  $T_N = 1501.3$  K. The reference for  $\Delta T_b$  is the first calibration on 10 November 2011 performed with HATPRO-G2 at the Jülich Research Center. (b) Repeatability of the LN<sub>2</sub> calibration. The mean difference  $\Delta T_b$  from the reference calibration (first calibration on November 10) and standard deviation  $\sigma(T_b)$  (error bars) of 11 LN<sub>2</sub> calibrations is given for all HATPRO-G2 channels. The colored numbers give the standard deviations  $\sigma(T_N)$  and  $\sigma(T_b)$ .

by assuming a maximum uncertainty of  $\pm 0.2$  K for the in-situ temperature measurement. This value is reasonable as it reflects the maximum deviation between the two temperature sensors within the ambient target during the campaign. For the three saturated V-band channels the resulting  $T_b$  uncertainty lies within  $\pm 0.2$  and  $\pm 0.3$  K. For smaller  $T_b$  values the impact of the uncertainty within the in-situ measurement drops linearly. All other channels are affected by approximately  $\pm 0.1$  K.

#### 4.1.5 Repeatability and validity

The repeatability is the capability of a calibration to reproduce the calibration parameters on time scales with negligible instrument drift. It is particularly important to assess the repeatability, because in this study a single LN<sub>2</sub> calibration is compared to a number of tipping curve calibrations. We address this issue by analyzing 11 LN<sub>2</sub> calibrations that were performed within about two hours, using the same radiometer that was deployed at RHUBC-II. The measurements were conducted at the Jülich Research Center on 10 November 2011. It is assumed that the repeatability is characterized by the stability of the noise diode temperature  $T_N$ , which is determined with every calibration. Using  $T_N$  from the 11 calibrations allows to map the uncertainty within  $T_N$  into the range of measured brightness temperatures. For these simulations the detector voltages, the target temperatures, and the other three calibration parameters in a 4-point calibration scheme (the system noise temperature  $T_R$ , detector gain  $g$  and the detector non-linearity parameter  $\alpha$ ) are kept constant while only  $T_N$  is varied.

The impact on measured  $T_b$  is demonstrated for the HATPRO-G2 channels at 51.26 GHz (Fig. 6). For this channel the standard deviation of  $T_N$  is  $\sigma(T_N) = 2.0$  K. Variations in  $T_N$  affect  $T_b$  measurements below the hot calibration point  $T_H$ . The repeatability can be expressed by the standard

deviation  $\sigma(T_b)$ .  $\sigma(T_b)$  does not only depend on the variation of  $\sigma(T_N)$ , but also on the distance between  $T_N$  and  $T_H$ , and the  $T_b$  value itself. The lower  $T_b$ , the larger the impact of  $T_N$  variations. For typical  $T_b$  values at 51.26 GHz under RHUBC-II conditions ( $T_b \approx 38.4$  K)  $\sigma(T_b)$  is 0.4 K. The results are summarized for all channels in Fig. 6. It shows that the repeatability  $\sigma(T_b)$  ranges between 0.2 and 0.4 K for the non-opaque channels. For the opaque channels above 54 GHz, the impact of  $T_N$  on  $T_b$  is negligibly small, because the calibration is dominated by the ambient target temperature. It is assumed that the uncertainty is mainly caused by noise and small variations from calibration to calibration in the level of LN<sub>2</sub> in the cold calibration load, which changes the resonance condition for standing waves between the receiver and the LN<sub>2</sub> surface.

The second point we would like to address here is the validity period of the LN<sub>2</sub> calibration. This aspect is rather important for the V-band channels which cannot be calibrated by the tipping curve calibration. Measurements of these channels depend on the noise diode temperature derived from a liquid nitrogen calibration. Our paper includes a comparison of the liquid nitrogen calibration and the tipping curve calibration. Results of the liquid nitrogen calibration, performed on 11 August 2009, are compared to tipping curve results from 16 August 2009. This comparison is only possible because the instrument's drift over this time period is negligible.

Like the repeatability, the drift is recorded by the variation of  $T_N$ . We have analyzed 27 liquid nitrogen calibrations performed with HATPRO-G2 at the Jülich Research Center between June 2010 and May 2012. The time period between subsequent calibrations ranges from less than one hour to several months. Still, a trend analysis of these calibrations reveals significant trends of  $T_N$ . Depending on the channel  $T_N$  changes by  $+0.006$  to  $+0.010$  K per day in the K-band

and by +0.054 to +0.072 K per day in the V-band (Table 3). The trends were calculated for a significance level of 0.05. The calculated trends are statistically significant for all channels apart from the channels at 31.40 GHz. The resulting confidence intervals vary between ±0.003 K and ±0.010 K per day (Table 3). On the basis of the calculated trends,  $T_N$  changes by less than +0.1 K in the K-band and +0.4 K in the V-band between 11 and 16 August 2009. The impact on  $T_b$  can be derived similarly to the repeatability analysis discussed above. We use the first calibration on 10 November 2011 performed in Jülich as reference. When this calibration is simulated with a  $T_N$  modified by the five-day drift,  $T_b$  is affected by less than  $0.06 \pm 0.01$  K for all channels. The given uncertainty results from the maximum uncertainty of the calculated trends. We can conclude that the  $T_N$  drift is negligible for the analysis in this study (Table 3).

### 4.2 Tipping curve calibration

The tipping curve calibration results and the associated uncertainties are investigated by a procedure that uses the elevation scans in the 70° N/250° N azimuthal plane, which were continuously performed during RHUBC-II. The results from these elevation scans cannot be used to determine the detector non-linearity  $\alpha$  and the noise diode temperature  $T_N$ , because no noise is injected during the scans. Nevertheless, calibrated zenith  $T_b$  values, which can be derived from each single scan, do not depend on these parameters. Measurements at elevations 90.0, 45.0, 30.0 and 15.0° are considered in the calibration procedure used for this analysis. Expressed in terms of air masses, they correspond to 1.0, 1.4, 2.0 and 3.0, respectively. This is an advantage compared to the method used by Han and Westwater (2000), who combine zenith measurements with only one additional measurement at an air mass that varies between 1.5 and 4. Fitting more than two opacity–air mass pairs allows the user to set quality thresholds that reduce the impact of atmospheric inhomogeneities on the calibration results (Sect. 4.2.5). From each scan passing the quality thresholds, zenith  $T_b$  is calculated from Eqs. (1), (8) and (9). Calibrated zenith  $T_b$  could then be used as a cold calibration point  $T_C$  to derive a  $T_b$  value for the scene. The impacts of different possible sources of uncertainty are examined in the following. Furthermore, the most appropriate quality thresholds are determined.

#### 4.2.1 Mean radiative temperature

The frequency-dependent mean radiative temperature of the atmosphere,  $T_{mr}$ , is a required parameter for the calibration. If a radiosonde profile exists, it can be calculated using a radiative transfer model to compute the absorption coefficient and optical depth:

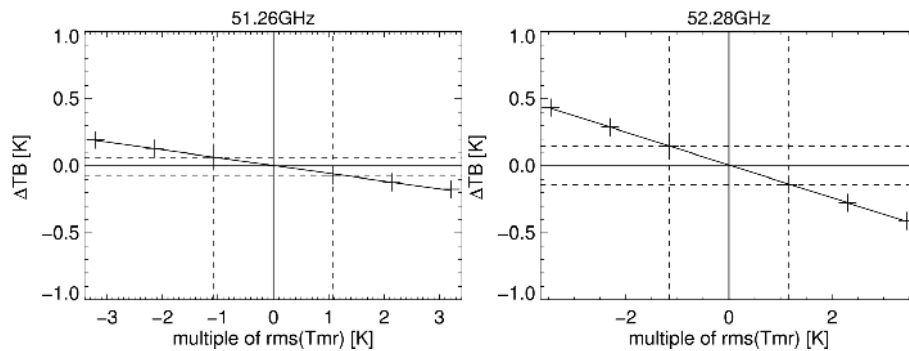
$$T_{mr} = \frac{\int_0^\infty T(s) \beta(s) e^{-\tau(0,s)} ds}{1 - e^{-\tau(0,\infty)}}, \quad (17)$$

**Table 3.** Results of 27 liquid nitrogen calibrations performed with HATPRO-G2 at the Jülich Research Center between June 2010 and May 2012.  $\overline{T_N}$  is the mean value of the 27 calibrations,  $T(T_N)$  is the calculated trend and its uncertainty (confidence interval, Sect. 4.1.5), D5 represents the estimated 5-day drift between 11 August 2009 (day of the LN<sub>2</sub> calibration for the measurement campaign RHUBC-II) and 16 August 2009 (day analyzed in this study), D74 estimates the drift over 74 days, corresponding to the duration of RHUBC-II.

$\nu$ [GHz]	$\overline{T_N}$ [K]	$T(T_N)$ [K day <sup>-1</sup> ]	D5( $T_b$ ) [K]	D74( $T_b$ ) [K]
22.24	401.1	+0.010 ± 0.003	-0.03	-0.51
23.04	399.0	+0.007 ± 0.002	-0.02	-0.37
23.84	353.4	+0.007 ± 0.002	-0.03	-0.42
25.44	342.5	+0.007 ± 0.002	-0.03	-0.42
26.24	370.5	+0.007 ± 0.002	-0.03	-0.42
27.84	363.3	+0.007 ± 0.002	-0.03	-0.41
31.40	340.1	+0.006 ± 0.010	-0.00	-0.04
51.26	1500.1	+0.054 ± 0.008	-0.05	-0.67
52.28	1341.8	+0.064 ± 0.010	-0.06	-0.82
53.86	1227.3	+0.066 ± 0.010	-0.04	-0.57
54.94	1181.3	+0.059 ± 0.010	-0.02	-0.23
56.66	1079.6	+0.061 ± 0.008	-0.01	-0.14
57.30	1083.8	+0.062 ± 0.008	-0.01	-0.13
58.00	1134.8	+0.072 ± 0.008	-0.01	-0.14

with the absorption coefficient  $\beta(s)$ , the physical temperature profile along the slant path,  $T(s)$ , and the total opacity  $\tau(0, \infty)$  along the slant path.  $T_{mr}$  must be determined prior to the calibration and cannot be determined from the radiometer itself. One possibility is to use a climatological  $T_{mr}$  derived from radiative transfer calculations with a climatology of atmospheric profiles. However, the tipping curve procedure is more accurate when the ambient surface temperature  $T_{surf}$  is used as predictor for  $T_{mr}$  in a linear regression scheme (Han and Westwater, 2000). Therefore,  $\tau$  is calculated from RHUBC-II radiosonde profiles using the Rosenkranz’98 absorption model (Rosenkranz, 1998). Then, the predictand  $T_{mr}$  is calculated for the nine non-opaque HATPRO-G2 channels from 22.24–52.28 GHz and different elevation angles from Eq. (17). The surface temperature, taken from the campaign’s AWS, is averaged over a time period of 5 min before and 1 h after launch of each radiosonde. Finally, linear regression coefficients are calculated from the obtained  $T_{mr}$ – $T_{surf}$  pairs. For 22.24 GHz the RMSE (Root Mean Square Error) is 3.2 K at three air masses. The RMSE decreases towards higher frequencies and reaches 1.1 K/1.2 K at 51.26 GHz/52.28 GHz at three air masses. The regression allows us to estimate  $T_{mr}$  from the continuous ambient surface temperature measurements throughout the campaign.

Figure 7 shows the sensitivity of calibrated zenith  $T_b$  to uncertainties in  $T_{mr}$ . Assuming an uncertainty of  $T_{mr}$  of one RMSE affects  $T_b$  in the V-band by  $\approx 0.1$  K. In the K-band the effect has virtually no impact. In order to exclude the effect



**Fig. 7.** Sensitivity of calibrated  $T_b$  to uncertainties in the  $T_{mr}$  calculations. Crosses:  $\Delta T_b$  resulting from a  $T_{mr}$  varied by  $\pm 3 \text{RMSE}(T_{mr})$  ( $\text{RMSE} = \text{Root Mean Square Error}$ ), solid line: linear fit, dashed: uncertainty area of one  $\text{RMSE}(T_{mr})$  and the resulting uncertainty of  $T_b$ . The  $\text{RMSE}(T_{mr})$  is 1.3 K (51.26 GHz) and 1.4 K (52.28 GHz).

of biases within the radiative transfer calculations, these are compared to results from the Liebe'93 (Liebe et al., 1993) and the Liebe'87 (Liebe, 1987) models. At three air masses the different models result in a  $T_{mr}$  that agrees within 0.1 K (Table 5).  $T_b$  results are not affected by such small differences. We can therefore conclude that  $T_{mr}$  can be derived from surface temperatures with sufficient accuracy for the tipping curve calibration.

#### 4.2.2 Air mass correction

The tipping curve procedure uses scene observations to derive opacities at different air mass values. For a plane-parallel non-refracting atmosphere the air mass depends only on the elevation angle  $\theta$ . In that case, the relative air mass is given by  $a_0 = 1/\sin(\theta)$ . However, for real observations the relative air mass differs from  $a_0$  when observing off zenith. On the one hand, the air mass is reduced due to the Earth curvature. On the other hand, atmospheric refraction increases the air mass. The effects from refraction and curvature increase when approaching the horizon and affect the tipping curve results. Even though the curvature effect is predominant, refraction is not negligible. Here, both effects are considered by a spherical ray-tracing algorithm that uses refractive index profiles derived from RHUBC-II radiosondes. The refractive index  $n$  is calculated at each level using a bulk formula:

$$n = 1 + \left( c_1 \frac{p - e}{T} + c_2 \frac{e}{T} + c_3 \frac{e}{T^2} \right) \times 10^{-6}, \quad (18)$$

with the atmospheric temperature  $T$  [K], the pressure  $p$  [hPa] and the partial pressure of water vapor  $e$  [hPa]. Here, the best available set of coefficients, suggested by Rüeiger (2002), is used:

$$c_1 = 77.695, \quad c_2 = 71.97, \quad c_3 = 3.75406 \quad (\text{in K hPa}^{-1}). \quad (19)$$

The ray-tracing algorithm calculates the slant path lengths for each height layer. The slant path opacity  $\tau(\theta)$  for the given radiosonde profile is then integrated over the full path length.

Finally, the exact relative air mass value is given by

$$a = \frac{\tau(\theta)}{\tau^{\text{zen}}}. \quad (20)$$

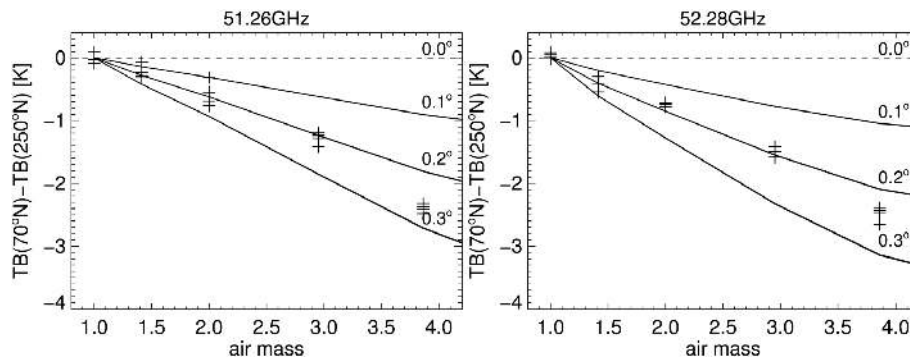
For each tipping curve scan,  $a$  is calculated by using the closest available radiosonde profile. The usage of  $a$  instead of  $a_0$  does not notably affect the tipping curve results for K-band channels under RHUBC-II conditions. For the two V-band channels, the impact is higher: zenith  $T_b$  results are affected by  $\Delta T_b^{\text{zen}} = +0.2$  K.

The uncertainty of the ray-tracing algorithm is assessed by the variation between all four radiosondes of the analyzed day. It is found that the uncertainty does not have any effect on the tipping curve results. Using the air mass correction suggested by Han and Westwater (2000) leads to an underestimation of  $T_b$  results by 0.1 K at 51.26 and 52.28 GHz, while the K-band is not affected.

#### 4.2.3 Beam width correction

The tipping curve calibration assumes that  $T_b$  measurements are given at distinct nominal air masses. However, microwave radiometers receive the incoming radiation over an aperture that is defined by the antenna pattern (Han and Westwater, 2000). For HATPRO-G2 side lobe suppression is better than  $-30$  dB. Numerical simulations of the beam pattern reveal that 99.5 % of the signal is received within an angular range of two half-power beam widths (HPBW). It is concluded that contribution from outside this range can be neglected.

Still, the finite HPBW affects the effective air mass value because  $T_b$  does not scale with the elevation angle. For transparent radiometer channels this results in an overestimated  $T_b$  because a larger part of the signal is received from directions below the nominal air mass value. This effect on measured  $T_b$  increases with air mass. In order to derive  $T_b$  measurements for the nominal air mass values, the specified HPBW is used (Rose et al., 2005). The beam is modeled as a Gaussian-shaped lobe with a doubled HPBW being



**Fig. 8.** Lines: radiative transfer calculations at different air masses for four clear sky radiosondes launched at 09:27, 10:39, 12:03, and 13:44 LT on 16 August 2009. Air masses are modified in order to simulate a tilt  $t$  of the radiometer by 0.1, 0.2, 0.3°:  $\Delta T_b = T_b(\text{el} = \text{el} + t) - T_b(\text{el} = \text{el} - t)$ , ( $\text{el}$  = elevation). Crosses: mean HATPRO-G2 measurements during radiosonde ascents (5 min before and 1 h after each launch). Radiative transfer calculations corrected air mass values; the antenna beam width and exact band-pass filters are considered.

7° for the K-band and 5° for the V-band channels. It is resolved by equally distributed and weighted radiative transfer calculations around the nominal elevation angle with a resolution of 0.1° (air mass correction is included). The corrections are applied to all  $T_b$  measurements within tipping curve scans on 16 August 2009. For air masses between one and three, the correction is below 0.1 K for the K-band channels due to the small opacities under the low water vapor conditions during RHUBC-II. At 51.26 and 52.28 GHz, the correction is 0.1 and 0.2 K at two air masses, and +0.5 and -0.5 K at three air masses, respectively. Zenith  $T_b$  values from the tipping curve calibrations of these two channels are affected by 0.5 K when using air masses between one and three. Han and Westwater (2000) derive a theoretical  $T_b$  correction from a Gaussian-shaped antenna pattern and an effective scale height. Compared to the numerical calculations with radiosonde profiles used here, that approach leads to a much smaller beam width correction: the impact on calibrated  $T_b$  would be less than 0.1 K at three air masses for all channels.

#### 4.2.4 Pointing error

The tipping curve procedure is very sensitive to uncertainties in the radiometer elevation pointing. A mispointing of 1° can easily lead to a calibration error of several Kelvin (Han and Westwater, 2000). As mentioned before, the radiometer continuously performed elevation scans along the (70° N/250° N) azimuthal plane. The V-band channels at 51.26 and 52.28 GHz show almost constant  $T_b$  differences between the two scanning directions that reach 2–3 K at four air masses. Model simulations show that a large part of these differences are systematic and can be explained by a tilt of the instrument by 0.2° (Fig. 8). Note that the simulations consider the beam width and air masses in a spherical and refracting atmosphere (Sect. 4.2.2).

Han and Westwater (2000) point out that tilts can be balanced by averaging measurements of symmetric elevation angle prior to the tipping curve procedure. Nevertheless, the random effect from atmospheric inhomogeneities cannot be compensated by this method. The effect of inhomogeneities on  $T_b$  increases with air mass as water vapor variations close to the surface are observed. This is reflected in the stronger variance of up to 0.7 K at four air masses when successive scans are compared (Fig. 8). This suggests that – depending on the measurement site and weather conditions – only measurements down to three air masses should be used for the tipping curve calibration. Furthermore, the considered air mass range may additionally be constrained by site-dependent obstacles.

For the RHUBC-II deployment, the detected instrument tilt is not compensated by averaging, because the degree of homogeneity differs significantly towards both scanning directions. For the K-band channels,  $T_b$  measurements towards both azimuthal directions deviate from each other by up to 2 K at four air masses. The difference can be explained by drier air masses over the hillside of Cerro Toco (70° N) compared to the open plateau (250° N). Consequently, in the K-band the number of succeeding successful scans towards both directions is reduced significantly. Therefore, scans towards both directions are separately used after correcting the elevation angles by 0.2°. Finally, it is assumed that the correction results in a residual pointing uncertainty of 0.05°. This uncertainty has no effect on the K-band channel results, though it increases with channel opacity and results in a  $\pm 0.1$  K uncertainty of calibrated  $T_b$  in the V-band (Table 5).

#### 4.2.5 Quality control and results

Five days after the original LN<sub>2</sub> calibration, the second operational day during RHUBC-II (16 August 2009) offered ideal clear sky conditions for analyzing continuous tipping curve

**Table 4.** Difference ( $\Delta T_b$ ) of mean zenith  $T_b$  (standard deviation) [K] from successful tipping curve calibrations (TIP) on 16 August 2009 to  $T_b$  values based on the LN<sub>2</sub> calibration on 11 August 2009.  $\Delta T_b = T_b(\text{TIP}) - T_b(\text{LN}_2)$ .  $\Delta T_b$  is calculated for different quality control thresholds within the tipping curve calibration (Sect. 4.2.5).  $n$  is the number of averaged successful tipping curve calibrations on 16 August 2009. Results from calibrations including a beam width correction and an exact air mass calculation are given in column “corr.”. The uncorrected results are given in column “uncorr.”.

corr <sub>min</sub> = 0.9995/χ <sub>max</sub> <sup>2</sup> = 10 <sup>-5</sup>				
ν [GHz]	corr.	<i>n</i>	uncorr.	<i>n</i>
22.24	-1.3 (0.2)	49	-1.2 (0.2)	39
23.04	-0.7 (0.2)	50	-0.7 (0.2)	49
23.84	-0.5 (0.2)	69	-0.5 (0.2)	68
25.44	-0.5 (0.2)	125	-0.4 (0.2)	133
26.24	+1.0 (0.2)	85	+1.0 (0.2)	95
27.84	-1.3 (0.1)	67	-1.3 (0.1)	71
31.40	-0.5 (0.2)	101	-0.5 (0.2)	101
51.26	-0.4 (0.4)	120	-0.2 (0.5)	130
52.28	-0.7 (0.3)	216	-0.4 (0.3)	223

calibration. In addition to the four available radio soundings, detector voltages and  $T_b$  data from HATPRO-G2 were taken.

The quality of a tipping curve scan is estimated by two criteria, which determine if the atmosphere is horizontally stratified to the required degree. Two measures are used to determine the goodness of the linear fit between measured opacities and air masses: the linear correlation coefficient  $\text{corr}(\tau, a)$  and  $\chi^2$ , which is sensitive to deviations of  $\tau$  from the linear relation to  $a$ . We are aware that these statistical quantities are not well defined for the small number of opacity–air mass pairs used. Nevertheless, they can be used to index the quality of the fit.  $\chi^2$  is defined in a relative way:

$$\chi^2 = \sum_{i=0}^n \frac{(\tau_i^{\text{calc}} - \tau_i^{\text{fit}})^2}{\tau_i^{\text{calc}}} \quad (21)$$

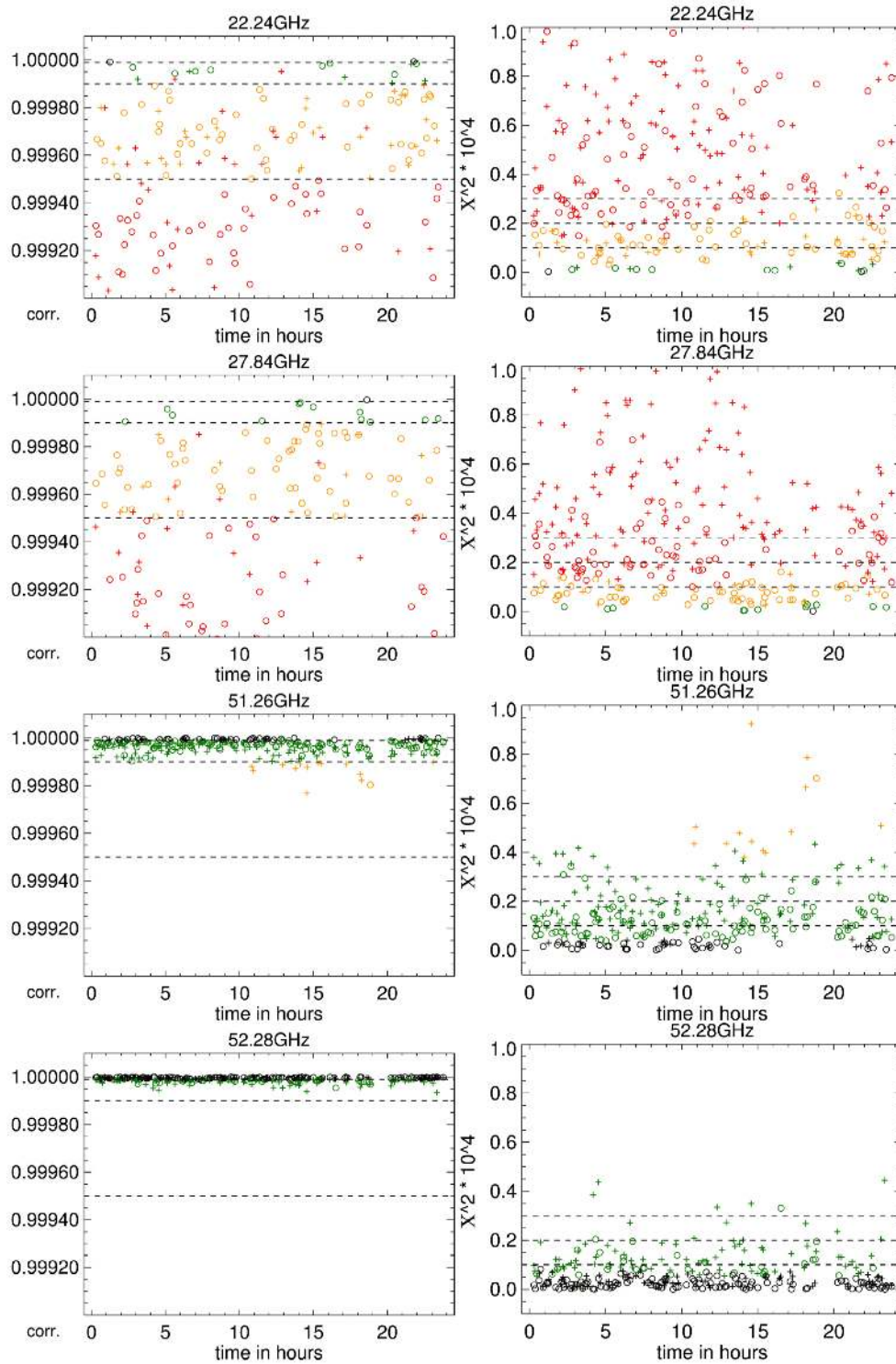
This study indicates that a relative threshold of  $\chi_{\text{max}}^2 = 1 \times 10^{-5}$  is most appropriate. Figure 9 illustrates the influence of different correlation thresholds on the tipping curve results. Naturally, raising the correlation threshold  $\text{corr}_{\text{min}}$  reduces the number of successful scans while increasing the calibration quality. Obviously, the quality threshold for  $\chi^2$  is a sensible supplement, because even scans with very high correlation coefficients fail. Generally, V-band channels show higher correlations than K-band channels, because the temperature field seems much more homogeneous than the atmospheric water vapor field. Different correlation thresholds have been tested and it is found that the manufacturer’s default threshold ( $\text{corr}_{\text{min}} = 0.9995$ ) is adequate for all calibrated channels. Higher values are not beneficial, because the standard deviation of all tipping curves on 16 August 2009 is

**Table 5.** Assessed total calibration uncertainties for RHUBC-II in Kelvin. For the LN<sub>2</sub> calibration, the total uncertainty results from uncertainties of the refractive index of the LN<sub>2</sub> surface ( $n_{\text{LN}_2}$ ), from resonances between the receiver and the LN<sub>2</sub> target (res), from uncertainties in the in-situ hot load measurement (hot), and from the detector non-linearity ( $\alpha$ ). The uncertainties are assessed for the mean  $T_b$  values measured at each channel on 16 August 2009. For the tipping curve calibration the total uncertainty results from the derivation of the mean radiative temperature ( $T_{\text{mr}}$ ), from the beam pointing (poi), and from atmospheric inhomogeneities (atm).

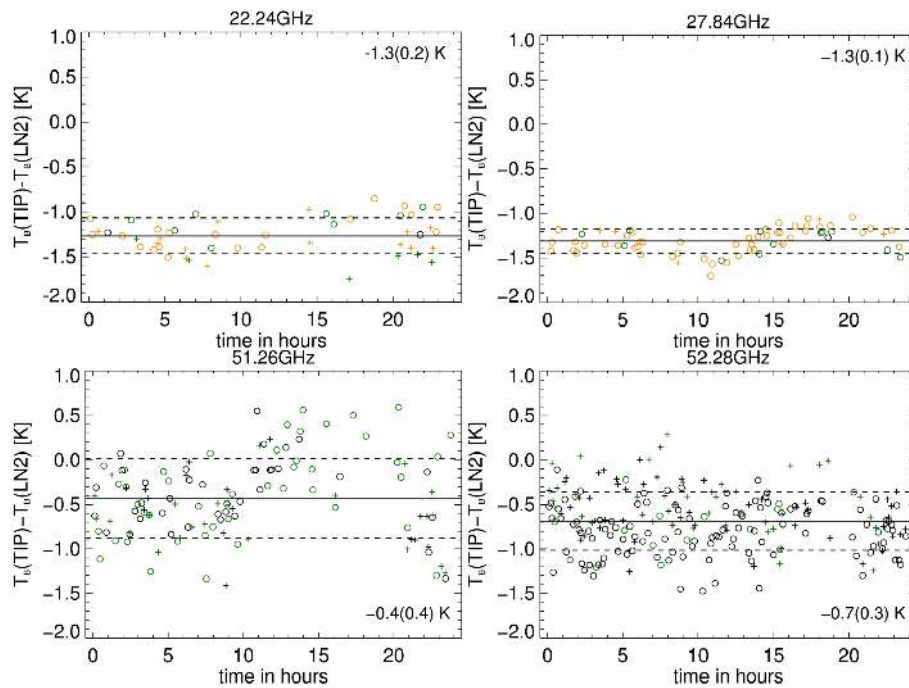
Liquid nitrogen calibration					
ν [GHz]	$n_{\text{LN}_2}$	res	hot	$\alpha$	total
22.24	±0.7	±0.4	±0.1	±0.04	±1.2
23.04	±0.7	±0.8	±0.1	±0.04	±1.6
23.84	±0.7	±0.2	±0.1	±0.03	±1.0
25.44	±0.7	±0.1	±0.1	±0.03	±0.9
26.24	±0.7	±0.3	±0.1	±0.03	±1.1
27.84	±0.7	±0.2	±0.1	±0.03	±1.0
31.40	±0.7	±0.2	±0.1	±0.02	±1.0
51.26	±0.6	±0.3	±0.1	±0.03	±1.0
52.28	±0.6	±0.1	±0.0	±0.00	±0.7
53.86	±0.4	±0.1	±0.0	±0.00	±0.5
54.94	±0.1	±0.0	±0.1	±0.00	±0.2
56.66	±0.1	±0.0	±0.2	±0.00	±0.3
57.30	±0.1	±0.0	±0.2	±0.01	±0.3
58.00	±0.1	±0.0	±0.1	±0.01	±0.2
Tipping curve calibration					
ν [GHz]	$T_{\text{mr}}$	poi	atm	total	
22.24	±0.0	±0.0	±0.2	±0.2	
23.04	±0.0	±0.0	±0.2	±0.2	
23.84	±0.0	±0.0	±0.2	±0.2	
25.44	±0.0	±0.0	±0.1	±0.1	
26.24	±0.0	±0.0	±0.2	±0.2	
27.84	±0.0	±0.1	±0.1	±0.1	
31.40	±0.0	±0.0	±0.2	±0.2	
51.26	±0.1	±0.1	±0.4	±0.6	
52.28	±0.3	±0.1	±0.3	±0.7	

not changed significantly. The number of successful tipping curves depends on the channel frequency: while it is small for the three K-band channels next to the line center (49–69), there are 120 and 216 successful tipping curves for the V-band channels at 51.26 and 52.28 GHz, respectively (Table 4). Again, this indicates the high inhomogeneity of atmospheric water vapor, which is most pronounced close to the line center at 22.235 GHz.

Figure 10 shows time series of zenith  $T_b$  differences between successful tipping curve calibrations and the original LN<sub>2</sub> calibration for selected channels. In Table 4 the results are summarized for all calibrated channels. The assumed quality thresholds result in a good repeatability of the tipping curve calibration throughout the day. The standard deviation of  $\Delta T_b$  over all scans is 0.1–0.2 K for the K-band



**Fig. 9.** Linear opacity–air mass correlations and  $\chi^2$  values for tipping curves on 16 August 2009 for selected HATPRO-G2 channels. Crosses: scans towards  $70^\circ$  N, circles: scans towards  $250^\circ$  N, dashed lines: quality thresholds  $\text{corr}_{\min}$  and  $\chi^2 = 0.1 \times 10^{-4}$ , red: failed tipping curves for  $\text{corr} < \text{corr}_{\min} = 0.9995$ ; tipping curve calibrations from successful elevation scans when only the correlation threshold is applied: green:  $\text{corr}_{\min} = 0.9995$ , black:  $\text{corr}_{\min} = 0.9999$ .



**Fig. 10.** Time series of tipping curve results  $T_b(\text{TIP})$  (16 August 2009) minus the original LN<sub>2</sub> calibration  $T_b(\text{LN}_2)$  (11 August 2009) for selected HATPRO-G2 channels. The daily mean value (standard deviation) is displayed within each subfigure. Crosses: one-sided elevation scans towards 70° N, circles: scans towards 250° N, black lines: mean calibrated zenith  $T_b$  and standard deviation (dashed) for scans towards 250° N; quality control thresholds are set to  $\text{corr}_{\min} = 0.9995$ ,  $\chi_{\max}^2 = 1.0 \times 10^{-4}$ .

channels. For the V-band channels at 51.26 and 52.28 GHz it is 0.5 and 0.3 K. It is assumed that this variability is induced by random processes such as atmospheric turbulence and radiometer noise. Therefore, it can be used to estimate the random uncertainty of tipping curve calibrations. Twice the standard deviation gives the maximum uncertainty range of measured  $T_b$  from a single tipping curve calibration (Table 5). One possibility to eliminate the random variability is to take the daily mean average of the tipping curve results. A similar approach is taken by Liljegren (1999), who introduces a rolling dataset of calibration parameters to the calibration procedure of ARM radiometers.

## 5 Summary and discussion

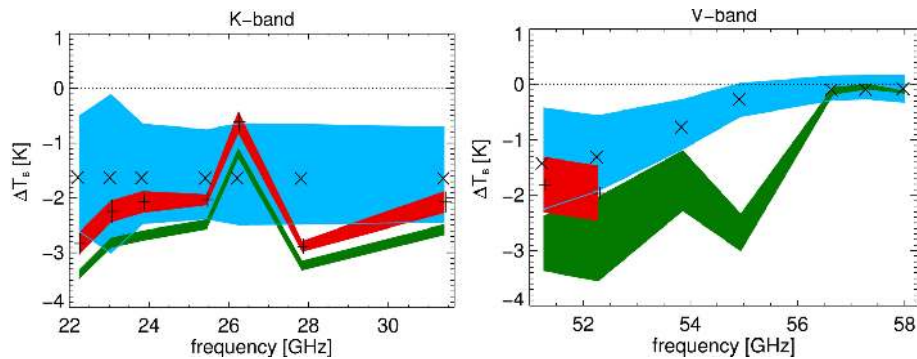
Two different calibration techniques for a state-of-the-art ground-based microwave radiometer (MWR) have been assessed on the basis of RHUBC-II data. The most common error sources that need to be considered when specifying the accuracy of the brightness temperature measurements are discussed. Table 5 gives a summary of the different uncertainty contributions. Although the uncertainties discussed with respect to the LN<sub>2</sub> calibration are partially instrument-specific, we note that RPG-HATPROs account for a major part of existing ground-based MWRs, and thus this analysis is valid for a significant part of the worldwide available

systems. The analysis of the tipping curve calibration is generally applicable to all ground-based MWRs.

### 5.1 Liquid nitrogen calibration

For the LN<sub>2</sub> calibration, a systematic error at 530 hPa was identified and corrected for by applying an improved formulation of the pressure-dependent boiling point of LN<sub>2</sub>. After this correction, the LN<sub>2</sub> calibration is mainly influenced by the temperature uncertainties of the cold and hot calibration targets. Here, the main contributions arise from the uncertainty of the reflective component that superposes the signal from the LN<sub>2</sub> target. The reflective component depends on the refractive index of the liquid nitrogen surface  $n_{\text{LN}_2}$  and the strength of the standing wave that builds up between the LN<sub>2</sub> surface and the receiver. This leads to  $T_b$  uncertainties, because  $n_{\text{LN}_2}$  is not exactly known and the development of the standing wave depends on the level of evaporating LN<sub>2</sub> in the cold load target. Resonances primarily affect the K-band channels, because here the largest amplitudes are observed. Additionally, both effects are frequency-dependent: the impact on calibrated  $T_b$  increases with decreasing  $T_b$ . Consequently, at RHUBC-II, K-band channels – with measured  $T_b$  below 10 K – are most affected.

Approaching the hot calibration point, uncertainties in the temperature measurement of the internal ambient target gain influence. It is shown that a non-zero reflectivity of the internal target can influence the derived  $T_b$  values. However, the



**Fig. 11.** Calibration comparison for HATPRO channels (left panel: K-band, right panel: V-band):  $\Delta T_b$  refers to the mean zenith  $T_b$  values based on the original LN<sub>2</sub> calibration on 11 August 2009 (boiling point correction Eq. 12). X: LN<sub>2</sub> calibration results with the improved boiling point correction (Eq. 15), light blue: estimated uncertainty range of the LN<sub>2</sub> calibration (Table 5), +: tipping curve results, red: estimated uncertainty range associated with the tipping curve calibration (Table 5), green: range of zenith  $T_b$  differences from radiative transfer calculations with three different absorption models: Liebe'87 (Liebe, 1987), Liebe'93 (Liebe et al., 1993) and Rosenkranz'98 (Rosenkranz, 1998). V-band calculations for < 56 GHz have been convolved with the real band-pass filter shapes measured with a resolution of 10 MHz.

effect is difficult to quantify because there is no information available on the uncertainty of the hot load reflectivity. In general, for different channels the overall uncertainty is between  $\pm 0.3$  and  $\pm 1.6$  K (Table 5). Furthermore, the LN<sub>2</sub> calibrations benefit from considering the detector non-linearity parameter  $\alpha$  within the 4-point calibration scheme. Compared to the 2-point calibration scheme, neglecting  $\alpha$  can result in a bias in the scene  $T_b$  in low-opacity channels that is close to 0.2 K.

## 5.2 Tipping curve calibration

Tipping curve results are very sensitive to the exact radiometer pointing. Therefore, scans on both sides of the radiometer were used to diagnose the instrument's tilt. The pointing is improved by compensating the instrument's tilt with a residual uncertainty of about  $0.05^\circ$ . The residual effect is negligible in the K-band, and even in the V-band  $T_b$  is only affected by 0.1 K. A further source of uncertainty is the derivation of the mean radiative temperature  $T_{mr}$ . The uncertainty of  $T_{mr}$  affects  $T_b$  by up to  $\pm 0.3$  K in the V-band. Finally, the uncertainty of successful tipping curve calibrations is dominated by atmospheric inhomogeneities that are not completely excluded by the quality thresholds (Sect. 4.2.5). The inhomogeneities lead to uncertainties up to  $\pm 0.2$  K in the K-band and up to  $\pm 0.7$  K in the V-band. The contribution from atmospheric inhomogeneities can be reduced by averaging the results of several tipping curve calibrations instead of the current practice to use a single tipping curve for an extended measurement period.

Compared to Han and Westwater (2000) more realistic air mass and beam width corrections – using observations from one to three air masses – are applied. In comparison to a plane-parallel non-refracting atmosphere, the used air mass correction that considers the Earth curvature and refraction (Rüeger, 2002) raises calibrated zenith  $T_b$  by 0.2 K

at 51.26 GHz and 0.1 K at 52.28 GHz. Neglecting the antenna beam width would increase zenith  $T_b$  measurements at RHUBC-II by up to 0.2 K.

In total, all effects add up to a zenith  $T_b$  uncertainty up to  $\pm 0.2$  K in the K-band and up to  $\pm 0.7$  K in the V-band. It is important to mention that these estimations depend on the atmospheric conditions at RHUBC-II. The assessed uncertainties may deviate for locations with higher opacities.

## 5.3 Comparative assessment

After assessing the uncertainties of the two calibration procedures, both are compared on the basis of calibrated zenith  $T_b$  on 16 August 2009. The average tipping curve results are compared to measurements based on the original LN<sub>2</sub> calibration five days before (Fig. 11). However, in Sect. 4.1.5 it is shown that the comparison is still possible, because the instrument drift over this period is negligible. Zenith  $T_b$  from tipping curves are generally smaller than  $T_b$  derived from the LN<sub>2</sub> calibration. Table 4 shows that for the chosen quality thresholds ( $\text{corr}_{\min} = 0.9995$ ,  $\chi_{\max} = 1 \times 10^{-5}$ ) the difference  $\Delta T_b = T_b(\text{TIP}) - T_b(\text{LN}_2)$  reaches the maximum of  $-1.3$  K.  $\Delta T_b$  does not seem to depend on the number of averaged tipping curves, because the repeatability of tipping curve results throughout the day is very good ( $< \pm 0.2$  K, Table 4). For the two V-band channels the difference is  $-1.8$  K at 51.26 GHz and  $-2.0$  K at 52.28 GHz. Additionally, Fig. 11 shows the results of the improved boiling point correction (Eq. 15). This correction significantly reduces the bias between the two calibration techniques to  $-0.4$  K at 51.26 GHz and  $-0.7$  K at 52.28 GHz (cf. Figs. 10 and 11). Apart from the channel at 27.84 GHz, the tipping curve calibration in the K-band generally provides smaller  $T_b$  results than the LN<sub>2</sub> calibration. Still, the total uncertainty ranges of both calibration methods overlap for the two V-band channels and for most K-band channels (Fig. 11) when the updated boiling point correction



is included. Only at 22.24 GHz and 27.84 GHz they are slightly outside the overlap range ( $-0.2$  and  $-0.4$  K, respectively). Nevertheless, the three lowest K-band channels show a bias, which systematically increases towards the line center. This raises the question whether an atmospheric water vapor signal has contaminated the LN<sub>2</sub> calibration via a reflective component (Sect. 4.1.2). Such an external atmospheric signal would reduce the reflective component and therefore calibrated  $T_b$ . An additional frequency-dependent effect results from the antenna's beam width, which increases with decreasing frequency. Therefore, the low-frequency channels could be influenced by an external signal that originates from objects at ambient temperature or from the atmosphere.

Finally, the question which technique gives results that are closer to the truth cannot be answered. However, radiative transfer calculations for the K-band channels and the results from the tipping curve calibration agree very well: Zenith  $T_b$  measurements that are calibrated by the tipping curve calibration only deviate by 0.3–0.8 K from simulated  $T_b$  values from RHUBC-II radiosonde profiles for all K-band channels (Fig. 11). Therefore, we recommend to use the tipping curve calibration for the K-band channels whenever possible (Table 5). This is especially true for low-opacity conditions, where the uncertainties of tipping curve results are small and the results show a very good repeatability (Fig. 10). In contrast, the LN<sub>2</sub> calibration suffers from large uncertainties at the cold calibration point. The influence of standing waves can be reduced by averaging calibration results. However, the  $T_b$  uncertainty resulting from an uncertain refractive index of LN<sub>2</sub> remains.

For the two V-band channels the uncertainty of the tipping curve is much larger than in the K-band, because most effects that impact the uncertainties grow with channel opacity. In the V-band, both calibration techniques are in good agreement when the LN<sub>2</sub> calibration is used with the improved boiling point correction. Also simulations for 51.26 and 52.28 GHz agree well with the tipping curve calibration results, though for the other non-opaque V-band channels below 56 GHz, simulated  $T_b$  are systematically smaller than  $T_b$  values from the improved LN<sub>2</sub> calibration. The difference reaches almost 3 K at maximum. Still,  $T_b$  uncertainties for non-opaque channels are smaller than for the LN<sub>2</sub> calibration, which suffers from resonances and an uncertain refractive index of LN<sub>2</sub> in the V-band, too. Of course, the LN<sub>2</sub> calibration results can be tuned to the tipping curve results by adapting  $n_{\text{LN}_2}$ . However, when doing this, each channel provides a significantly different  $n_{\text{LN}_2}$  value. Furthermore, this excludes an independent comparison of both calibration techniques.

## 6 Outlook

Summarizing, a larger number of detailed uncertainty assessments are needed to explain the differences between the LN<sub>2</sub> calibration and the tipping curve calibration. Upcoming measurement deployments should be accompanied by an accuracy analysis of the two calibration methods. Particularly, it would be valuable to repeat the sensitivity study for radiometer measurements under different atmospheric conditions, because RHUBC-II covers only the extreme case of dry and relatively warm conditions at high altitudes. Parts of the estimated calibration uncertainties might scale with  $T_{\text{mr}}(1 - \exp(-\tau))$  of the observed atmosphere, as suggested by one of the reviewers of this study. Other parts depend on the instrument type and the calibration procedure. At this point future deployments will help estimate how general the assessed calibration uncertainties are.

An important aspect concerning the LN<sub>2</sub> calibration is to verify its reproducibility by repeated calibrations. Additionally, frequent elevation scans help to improve the calibration accuracy by the tipping curve method. Using both methods simultaneously gives the opportunity to detect possible calibration problems.

Finally, when the calibration accuracy is properly assessed, the resulting  $T_b$  measurements can be used to validate gas absorption models. In contrast to the K-band, where different absorption models show a good agreement, in the V-band  $T_b$  calculations from different models deviate by up to 2 K. Any significant deviation from the measurement can then be assigned to the model spectroscopy, i.e., the absorption parameters. Here, RHUBC-II data give the opportunity to validate gas absorption models in a very dry middle-to-upper troposphere.

*Acknowledgements.* The data used in this paper were collected as part of the ARM Program sponsored by the US DOE, Office of Science, Office of Biological and Environmental Research, Climate and Environmental Sciences Division. The microwave radiometer HATPRO-G2 was funded by the German Science Foundation (DFG, INST 216/519-1FU66).

The experiments on HATPRO-G2's LN<sub>2</sub> calibration at the Jülich Research Center are also part of Nils Küchler's bachelor thesis. Thank you, Nils, for contributing the statistical analysis of the observed standing waves.

We gratefully acknowledge financial support by the SFB/TR 32 "Pattern in Soil–Vegetation–Atmosphere Systems: Monitoring, Modelling, and Data Assimilation" funded by the Deutsche Forschungsgemeinschaft (DFG).

Edited by: D. Feist

## References

- Benson, J., Fischer, J. and Boyd, D. A.: Submillimeter and Millimeter Optical Constants of Liquid Nitrogen, *Int. J. Infrared Milli.*, 4, 145–152, doi:10.1007/BF01008973, 1983.
- Cadeddu, M. P., Payne, V. H., Clough, S. A., Cady-Pereira, K., and Liljegren, J. C.: Effect of the Oxygen Line-Parameters Modeling on Temperature and Humidity Retrievals From Ground-Based Microwave Radiometers, *IEEE T. Geosci. Remote*, 45, 2216–2223, doi:10.1109/TGRS.2007.894063, 2007.
- Güldner, J. and Spänkuch, D.: Remote Sensing of the Thermodynamic State of the Atmospheric Boundary Layer by Ground-Based Microwave Radiometry, *J. Atmos. Ocean. Tech.*, 18, 925–933, doi:10.1175/1520-0426(2001)018<0925:RSOTTS>2.0.CO;2, 2000.
- Han, Y. and Westwater, E. R.: Analysis and Improvement of Tipping Calibration for Ground-Based Microwave Radiometers, *IEEE T. Geosci. Remote*, 38, 1260–1276, doi:10.1109/36.843018, 2000.
- Hewison, T. J.: 1D-VAR Retrieval of Temperature and Humidity Profiles from a Ground-based Microwave Radiometer, *IEEE T. Geosci. Remote*, 45, 2163–2168, doi:10.1109/TGRS.2007.898091, 2007.
- Hewison, T. J., Cimini, D., Martin, L., Gaffard, C., and Nash, J.: Validating clear air absorption models using ground-based microwave radiometers and vice-versa, *Meteorol. Z.*, 15, 27–36, doi:10.1127/0941-2948/2006/0097, 2006.
- Hosking, M. W., Tonkin, B. A., Proykova, Y. G., Hewitt, A., McAlford, N. N., and Button, T. W.: The Dielectric Constant of Liquid Nitrogen over the Frequency Range 0.5 to 10.4 GHz, *Supercond. Sci. Technol.*, 6, 549–552, doi:10.1088/0953-2048/6/7/018, 1993.
- Liebe, H. J. and Layton, D. H.: Millimeter-wave properties of the atmosphere: Laboratory studies and propagation modeling, *Natl. Telecommun. and Inf. Admin.*, Boulder, Colorado, 1987.
- Liebe, H. J., Hufford, G. A., and Cotton, M. G.: Propagation Modeling of Moist Air and Suspended Water/Ice Particles at Frequencies below 1000 GHz, *Atmospheric Propagation Effects through Natural and Man-Made Obscurants for Visible to mm-Wave Radiation*, AGARD Conference, Palma de Mallorca, Spain, 1993.
- Liljegren, J. C.: Automatic Self-Calibration of ARM Microwave Radiometers, Ninth ARM Science Team proceedings, 22–26 March 1999, San Antonio, Texas, 1999.
- Liljegren, J. C.: Microwave Radiometer Profiler Handbook, Manual, Argonne National Laboratory, Washington, D.C., 2002.
- Löhnert, U. and Maier, O.: Operational profiling of temperature using ground-based microwave radiometry at Payerne: prospects and challenges, *Atmos. Meas. Tech.*, 5, 1121–1134, doi:10.5194/amt-5-1121-2012, 2012.
- McGrath, A. and Hewison, T. J.: Measuring the Accuracy of MARSS – An Airborne Microwave Radiometer, *J. Atmos. Ocean. Tech.*, 18, 2003–2012, doi:10.1175/1520-0426(2001)018<2003:MTAOMA>2.0.CO;2, 2001.
- NIST – National Institute of Standards and Technology: <http://webbook.nist.gov> (last access: 1 October 2013), 2013.
- Radiometer Physics GmbH: Instrument Operation and Software Guide – Principle of Operation and Software Description for RPG standard single-polarization radiometers, available at: <http://www.radiometer-physics.de/rpg/html/Download.html>, last access: 14 June 2013.
- Radiometrics: Profiler Operator’s Manual for the MP-3000A, available from: [Info@radiometrics.com](mailto:Info@radiometrics.com), 2007.
- Randa, J., Walker, D. K., Cox, A. E., and Billinger, R. L.: Errors resulting from the reflectivity of calibration targets, *IEEE T. Geosci. Remote*, 43, 50–58, doi:10.1109/TGRS.2004.839809, 2005.
- Reesor, G. E., Dagg, I. R., and Mohabi, S.: The Complex Dielectric Constant of Liquid Nitrogen in the Region 18 to 26 GHz, *Can. J. Phys.*, 53, 2611–2612, doi:10.1139/p75-318, 1975.
- Rose, T., Crewell, S., Löhnert, U., and Simmer, C.: A network suitable microwave radiometer for operational monitoring of the cloudy atmosphere, *Atmos. Res.*, 75, 183–200, doi:10.1016/j.atmosres.2004.12.005, 2005.
- Rosenkranz, P. W.: Water vapor microwave continuum absorption: A comparison of measurements and models, *Radio Sci.*, 33, 919–928, doi:10.1029/98RS01182, 1998.
- Rüeger, J. M., *Refractive Index Formulae for Radio Waves*, JS83 Integration of Techniques and Corrections to Achieve Accurate Engineering, FIG XXII International Congress, Washington, 2002.
- Rutland, J.: On the extreme aridity of coastal and Atacama deserts in Chile, Ph.D. thesis, University of Wisconsin, Madison, 1977.
- Shitov, S. V., Uvarov, A. V., Ermakov, A. B., and Vystavkin, A. N.: Measurement of the Reflector Antenna Absorption at Terahertz Frequencies by the Method of a Balanced Radiometer, *Tech. Phys.*, 56, 1652–1660, doi:10.1134/S1063784211110223, 2011.
- Span, R., Lemmon, E. W., Jacobsen, R. T., Wagner, W., and Yokozeki, A.: A Reference Equation of the State for the Thermodynamic Properties of Nitrogen for Temperatures from 63.151 to 1000 K and Pressures to 2200 MPa, *J. Phys. Chem. Ref. Data*, 29, 1361–1433, doi:10.1063/1.1349047, 2000.
- Stokes, G. M. and Schwartz, S. E.: The Atmospheric Radiation Measurement (ARM) Program: Programmatic Background and Design of the Cloud and Radiation Test Bed, *B. Am. Meteorol. Soc.*, 75, 1201–1221, doi:10.1175/1520-0477(1994)075<1201:TARMPP>2.0.CO;2, 1994.
- Turner, D. D. and Mlawer, E. J.: The Radiative Heating in Underexplored Bands Campaigns, *B. Am. Meteorol. Soc.*, 91, 911–923, doi:10.1175/2010BAMS2904.1, 2010.
- Turner, D. D., Mlawer, E. J., Bianchini, G., Cadeddu, M. P., Crewell, S., Delamere, J. S., Knuteson, R. O., Maschwitz, G., Mlynz-cak, M., Paine, S., Palchetti, L., and Tobin, D. C.: Ground-based high spectral resolution observations of the entire terrestrial spectrum under extremely dry conditions, *Geophys. Res. Lett.*, 39, L10801, doi:10.1029/2012GL051542, 2012.
- Vinogradov, E. A., Dianov, E. M., and Irisova, N. A.: Dielectric characteristics of liquid nitrogen at  $\lambda = 2.3$  mm, *Sov. Phys. Tech. Phys.*, 11, p. 983, 1967.
- Westwater, E. R.: Ground-based passive probing using the microwave spectrum of oxygen, Troposphere kinetic temperature structure from ground-based measurements of oxygen emission spectra in microwave region, using inversion technique and iterative method, *J. Res. Natl. Inst. Stand. Technol.*, 69D, 1201–1211, 1965.