Is There Really a Closure Gap Between 183.31-GHz Satellite Passive Microwave and *In Situ* Radiosonde Water Vapor Measurements?

Oleksandr Bobryshev, Stefan A. Buehler, Viju O. John, Manfred Brath, and Hélène Brogniez

***Abstract*— We present a new closure study between radiosonde and microwave satellite humidity measurements. The radiosonde data are from the Global Climate Observing System Reference Upper-Air Network. The satellite data are from the radiometers: MHS, Advanced Technology Microwave Sounder, and Sondeur Atmosphérique du Profil d'Humidité Intertropicale par Radiométrie. Like previous studies, we find the satellite data to be "colder" than simulated radiosonde data. But the mean bias value (0.4 K) is smaller than previously reported and, according to our analysis, not significant. The error budget suggests an uncertainty of 0.52–1.06 K. We also show that the improvement in closure can be attributed to improvements in the intercomparison methodology.**

***Index Terms*— Humidity measurement, microwave measurements, remote sensing.**

I. Introduction

OPERATIONAL satellite measurements at the water vapor (WV) absorption line centered at 183.31 GHz (hereinafter referred as the WV-line) started in 1991 with the advent of the Special Sensor Microwave/Temperature-2 radiometer onboard the Defense Meteorological Satellite Program (DMSP-F11) satellite [1]. Currently, there are several sensors which have channels at the WV-line, such as Advanced Microwave Sounding Unit-B (AMSU-B), Microwave Humidity Sounding, Special Sensor Microwave Imager/Sounder, and Sondeur Atmosphérique du Profil d'Humidité Intertropicale par Radiométrie (SAPHIR). There will be a continuation of these measurements at least up to 2043 with programs like the Joint Polar Satellite System Advanced Technology Microwave Sounder (ATMS) radiometer, the European Microwave Sounder radiometer onboard Metop Second Generation (Metop-SG-A) and the Ice Cloud Imager radiometer onboard Metop-SG-B, or the Chinese Microwave Humidity Sounder (MWHS-3) radiometer. These radiometers measure the intensity of the radiation in different frequencies around the center of the WV-line, which allows the WV profile to be derived [2].

The data from these instruments are often compared with *in situ* measurements for quality evaluation. Recently, two independent studies, in [3] and [4], have found a significant disagreement between radiosonde *in situ* measurements and the satellite measurements. On average, the satellite data appeared to be "colder" than the radiosonde-based simulated satellite data (the difference expressed in the brightness temperature T_B). These studies also found the disagreement to increase with the distance of the channel position from the WV-line center. As discussed in [5], it seems unlikely that this behavior is simply due to miscalibration of the radiometers or the radiosondes. Hence, this closure gap, this disagreement between the different methods to measure tropospheric humidity, questions our fundamental understanding of the atmosphere and the measurement systems and the way we compare them.

This paper is a follow-up to the works of [3] and [4]. We aim to check the existence of the systematic difference between radiosonde and satellite data, the result found in the above-mentioned papers. We use a longer up-to-date data set and put special focus on the setup of the comparison.

In situ radiosonde and remote satellite measurements cannot be compared directly because of their different natures and the different measurement scales. One of the common methods to compare these two measurements is to use the radiosonde profile as an input for a radiative transfer model (RTM) and then compare the output of the RTM with the satellite data, both expressed in units of brightness temperature (T_B) [6]. Another aspect of this comparison is that the radiosonde provides *in situ* measurements of the troposphere, and the obtained vertical profile of the atmosphere is treated as a point measurement at the certain location. On the other hand, satellites provide area measurements, one MHS pixel at nadir representing a disk of 16 km in diameter. The problem of such a comparison is the spatial scale of the measurements. Small-scale inhomogeneities, on the kiolometer scale and smaller, can influence the measurement, and there is no way to compensate for these effects. For example, the satellite may sense cloudy sky, but the radiosonde may ascend through the clear part of the scene. The different durations of the measurements,

Manuscript received September 4, 2017; revised November 2, 2017; accepted December 14, 2017. *(Corresponding author: Oleksandr Bobryshev.)*

O. Bobryshev, S. A. Buehler, and M. Brath are with the Meteorological Institute, Center for Earth System Research and Sustainability (CEN), Department of Earth Sciences, Universität Hamburg, 20146 Hamburg, Germany (e-mail: oleksandr.bobryshev@uni-hamburg.de).

V. O. John is with the European Organisation for the Exploitation of Meteorological Satellites, 64295 Darmstadt, Germany, and also with the Met Office Hadley Centre, Exeter EX1 3PB, U.K.

H. Brogniez is with the University of Versailles Saint Quentin-en-Yvelines, 78000 Versailles, France.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2017.2786548

parts of a second for satellite, and hours for radiosonde also add complexity to the comparison. The common approach, also used in this paper, is to look for cases when satellite and radiosonde measurements occurred sufficiently close in time, so-called collocations or matchups, so that they measured the "same" atmosphere. Normally several hundred to several thousand matchups are used in this type of comparison. The remaining paper is organized as follows: the methodology and the data sets used are described in Section II. Section III presents our results, as well as a discussion of the differences to those from [3] and [4]. We also present the error model of the comparison with assessment of individual contributions. In Section IV, we draw up the conclusions of this paper.

II. Method and Data

A. Methodology

As described in Section I, we feed radiosonde data into an RTM to generate simulated satellite observations. From each radiosonde record, we use the vertical profiles of pressure, temperature, altitude, and WV mixing ratio [6]. Reference [7] showed that neglecting O_3 data when doing simulations for the microwave region leads to a systematic error of up to 0.5 K for some frequencies. The analysis showed that monthly mean climatological ozone profiles are sufficient to account for the effect of these lines on AMSU-B radiances. Therefore, for all the radiosonde profiles, we add the climatological ozone profiles from the Fast Atmospheric Signature Code (FASCODE) atmospheric data set [8].

Before comparing the two measurements, we select the cases when radiosonde and satellite measurements occurred sufficiently close in time and space. We require the difference in time between the satellite overpass and the radiosonde launching time to be within ± 3 h. We use two filters to insure the stability of the WV field: radiosonde horizontal drift during the ascent and the weighting of the individual collocations based on the inverse of the standard deviation squared of the satellite T_B inside the "target area." These are described in detail later in this section. The next complication is how to compare point measurements (radiosondes) with area measurements (satellite). We compare one radiosonde-simulated T_B with the averaged satellite T_B obtained in the "target area." The target area is a circle with 50-km radius around the radiosonde launching point, which contain all the satellite pixels (in this paper, it is 5–20 pixels). We calculate the radiosonde-simulated T_B for the average viewing angle of all satellite pixels inside the "target area" to account for the limb effect in the satellite data. This methodology was first suggested and described in detail in [6].

In this paper, we introduce additional steps to the methodology that are unique for this paper. The first step is the filtering of the radiosonde data in order to reduce the variability caused by the horizontal inhomogeneity of the atmosphere. We select only radiosonde profiles with a drift from the launching position less than 15 km based on the GPS-data. The second step is to filter all the matchups where one or both of the measurements, the satellite or the radiosonde, include a cloud signal. For this purpose, we use the Cloud Mask Product (CMA) data from the Spinning Enhanced Visible Infra-Red Imager (SEVIRI) instrument onboard the Meteosat Second Generation (MSG) geostationary satellite [9]. The CMA is based on the infrared (IR) measurements which are more sensitive than microwave-based cloud detection. Still in some cases, low- and middle-level clouds can be undetected [10]. The third step is to give higher weight to the matchups that we have higher confidence in. The measure for the weighting is the inverse of the standard deviation squared of the satellite T_B inside the "target area." We assign specific weight to each specific matchup. A high standard deviation corresponds to the heterogeneous state of the "target area," whereas low standard deviation corresponds to the homogeneous scene, with similar satellite T_B. The matchups with smaller standard deviation have higher weight, when computing the mean value of the bias. This step was already introduced in [6, Sec. 3.5], but we put this step among the additional steps to highlight its importance. The fourth step is to use more accurate radiosonde input data from the reference quality Global Climate Observing System Reference Upper-Air Network (GRUAN) radiosonde network and a bigger number of satellite data. The input data as well as the RTM, used to simulate satellite measurements, are presented in detail in Section II-B.

Major Differences From Previous Studies: Two earlier studies [3], [4] came to the conclusion that there is indeed a closure gap between the radiosondes and the satellites. These studies were set up quite similar to ours, in that they identify matchups between the radiosonde launch and a "target area" in the satellite data. Like us, they used some cloud filtering. Major differences between this paper and these previous studies are utilization of different radiosounding data, cloud identification techniques, and the methodology for selecting the "valid" matchups for the analysis. First, we use the GRUAN-processed radiosonde data, whereas they used a mix of the Vaisala processed and the National Center for Atmospheric Research (NCAR) corrected radiosonde data (the NCAR correction is the same as the GRUAN correction). Second, we do the cloud identification based on IR satellite data, as described in Section II-A, whereas they did cloud identification based on the radiosonde relative humidity measurement and T_B difference between microwave channels. Third, in contrast to the earlier studies, we weight individual matchups based on the inverse of the standard deviation squared of the T_B among the pixels in the "target area." We examine in detail the effect of these differences in Section III.

B. Data

The GRUAN operates under the joint governance of Global Climate Observing System and the World Meteorological Organization [11]. It aims to be the ground-based network for reference observations of upper-air climate parameters. There are strict quality criteria for the selected stations. Every station must use the same pre- and postlaunch controls, use the same type of radiosondes, provide extensive and complete metadata for each launch, use well-documented correction algorithms, and have vertically resolved measurement uncertainties. In this paper, we utilize the "RS92 GRUAN Data Product (Version 2)." We consider data from five stations: two in the

TABLE I

SUMMARY OF RADIOSOUNDING STATIONS USED IN THIS PAPER: LOCATIONS OF THE STATION, NUMBER OF SATELLITE-RADIOSONDE MATCHUPS, AND TIME COVERAGE WHEN THEY OCCUR. ALL THE DATA PRESENTED AFTER ALL FILTERS ARE APPLIED. THE GRUAN-PROCESSED RADIOSONDE DATA FOR THE TROPICAL STATIONS MANUS AND NAURU ARE AVAILABLE ONLY FOR YEARS 2011–2014 AND 2011–2013, RESPECTIVELY

Station	Location		# of collocations & time range					
	Lat	Lon	MHS		ATMS		SAPHIR	
Lindenberg, Germany	52.21^0	14.12^0	1575	(2009-2015)	369	(2012-2015)		
Lamont, OK, USA	36.60^0	-97.49^0	824	(2009-2015)	199	(2012-2015)		
Manus	-2.06^0	147.42^0	232	(2011-2014)	45	(2011-2014)	84	(2011-2014)
Nauru	-0.52^0	166.92^0	101	(2011-2013)	23	(2011-2013)	56	(2011-2013)
Barrow, AK, USA	71.32^0	-156.61^0	273	(2009-2015)	70	(2012-2015)		

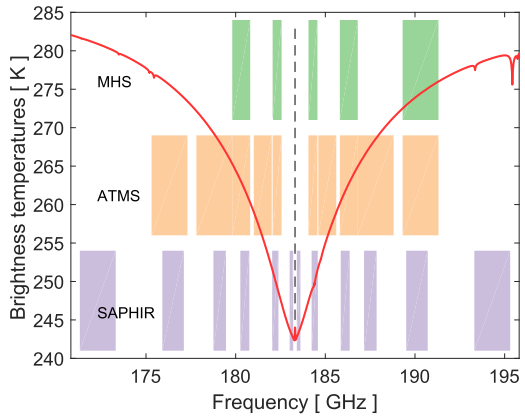


Fig. 1. The brightness temperatures simulated for the frequency range centered around the WV absorption line centered at the 183.31 GHz (black dotted line). FASCODE midlatitude summer atmosphere [8] served the input profile for the simulation. The boxes show the width and positions of the channels of passive microwave sensors MHS (green), ATMS (orange), and SAPHIR (violet).

tropics, two in the midlatitudes, and one in the subpolar region. The exact position of the radiosonde stations, the number of spatial-temporal radiosonde-satellite collocations used in this paper, and the time range are summarized in Table I.

Among the satellite measurements in this paper, we utilize data from three types of radiometer, with sounding channels at the WV-line. An overview of all the instrument channels as well as the absorption line is presented in Fig. 1. The following radiometers and products were used in this paper: MHS level 1B data (raw counts format) converted using the AAPPv7 program [12] to level 1C (T_B); SAPHIR level 1A2 data (T_B); and ATMS Sensor Data Record (ATMS-SDR, T_B) data. We took the MHS and ATMS data from the National Oceanic and Atmospheric Administration (NOAA) Comprehensive Large Array-Data Stewardship System (CLASS) Archive, and we took the SAPHIR data from the ICARE Data and Services Center.

The first radiometer is the Microwave Humidity Sounder (MHS). It is mounted on four operational meteorological satellites. We use the data from three of them: the polar-orbiting USA satellite NOAA-18 and the European satellites Meteorological operational satellite (MetOp-A) and MetOp-B [European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT)]. MHS samples in three channels near the WV-line (183 ± 1.0, 183 ± 3.0, *and* 190 GHz) have a swath width of approximately 2300 km with a scan angle of $\pm 49.44°$ around nadir, sampling 90 contiguous pixels. This scan pattern and geometric resolution translate to a 17-km diameter pixel at nadir [13]. MHS radiometer onboard NOAA-19 has a high level of noise at Channel 3 (183 ± 1.0 GHz) [14], thus we did not use the data from this instrument.

The second radiometer is the SAPHIR. It is mounted on the Megha-Tropiques satellite mission, which is operated jointly by the space agencies of India and France. SAPHIR samples in six channels near the WV-line (183 ± 0.2, 183 ± 1.1, 183 ± 2.7, 183 ± 4.2, 183 ± 6.6, and 183 ± 11.0 GHz) have a swath width of about 1700 km with a scan angle $\pm 42.96°$ around nadir, sampling 130 contiguous pixels. The pixel size at nadir is 10 km in diameter [15], [16].

The third instrument is the ATMS. It is mounted on the USA polar satellite Suomi National Polar-orbiting Partnership. ATMS samples in five channels near the WV-line (183 ± 1.0, 183 ± 1.8, 183 ± 3.0, 183 ± 4.5, and 183 ± 7.0 GHz) have a swath width of 2300 km with a scan angle of $\pm 52.73°$, and a 16 km in diameter pixel at nadir [17].

C. Radiative Transfer Model

The Atmospheric Radiative Transfer Simulator (ARTS) is a free, open-source software package, primarily designed for thermal radiative transfer simulations. It is a line-by-line RTM capable of simulating in any observational geometry and mimicking radiometer characteristics [18], [19]. Inputs to ARTS are the profiles of pressure, temperature, and absorbing species. We used the radiosonde measurements as input supplemented with climatological O_3 data. Absorbing species are H_2O, O_2, O_3, and N_2. Volume mixing ratios of O_2 and N_2 are taken constant at 0.2095 and 0.7808 throughout the profile. The line shape for all the absorbing species is Voigt–Kuntz [20] with van Vleck and Huber forefactor [21]. The spectral line parameters are taken from the High Resolution Transmission (HITRAN-2012) catalog, with one notable exception. For the WV-line, the self- and air broadened width parameters and the temperature exponents are taken according to the recommendations of [22]. Also, for H_2O, O_2, and N_2, we use the MT_CKD-252 continuum model [23]. The surface is treated specularly with a constant emissivity of 0.95. Clearsky conditions are assumed for all calculations. To simulate the real sensors, ARTS uses a sensor model for each radiometer with the precise characteristics, such as sideband frequencies [24]. ARTS performs a series of monochromatic calculations inside the defined bands; first lower and upper bands are averaged internally; then, to get the channel

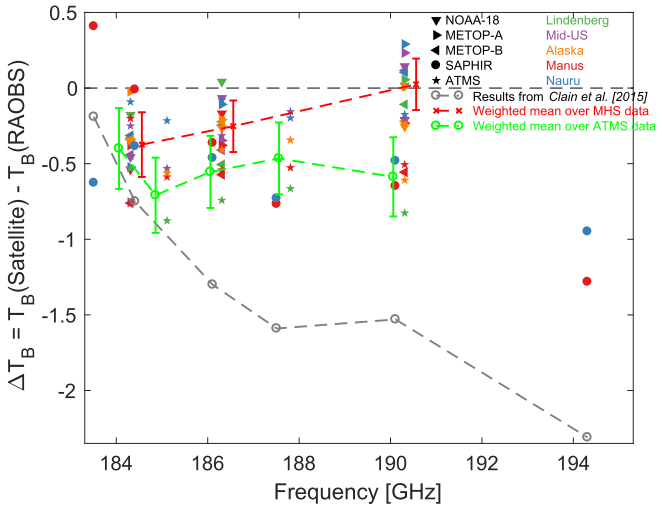


Fig. 2. Mean T_B difference between the satellite and the radiosonde on one side of the WV-line for all combinations of the radiosonde stations and satellite radiometers for different satellite channels. For simplicity, only one side of the absorption line is shown. Detailed description is in Section II-B. Specific radiometers are represented by the different marker types. The radiosounding stations are represented by different colors. The red and the green lines show the mean T_B values for all the satellite-radiosonde collocations of MHS and ATMS radiometers. The gray line presents the T_B difference as found in [3] for the day and night collocations, using 25 adjacent pixels (Table 4 in the publication). The exact values of the weighted mean, used to create this plot, and weighted standard deviation are given in the Supplementary Material.

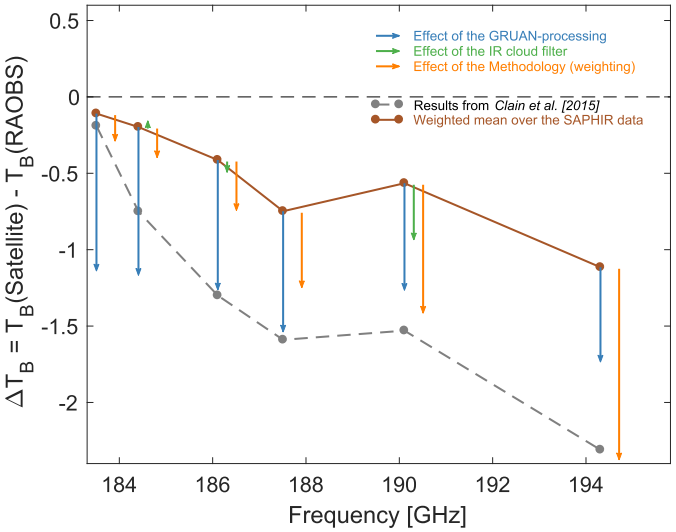


Fig. 3. Influence of the steps made differently in this paper compared with the studies of [3] and [4]: methodology of comparing the satellite measurements with radiosonde measurements, more specifically the influence of the radiosonde data processing GRUAN and Vaisala; the effect of the IR cloud filter, compared with the microwave-based cloud filters; assigning different weights to individual collocations according to the inverse of the standard deviation squared of the satellite T_B inside the "target area." The brown line is the mean value for all radiosounding stations for the SAPHIR radiometer. The gray dashed line presents the T_B difference as found in [3], the same as Fig. 2. Radiosonde-satellite pair data used to estimate the effects: 1) of the GRUAN-processing: Manus and SAPHIR; 2) the IR cloud filter: Lindenberg, MHS, and ATMS (these radiometers do not sample at the frequencies of SAPHIR channels 1, 3 and 6, thus no estimate of the IR cloud filter effect); 3) the methodology: Manus, Nauru, and SAPHIR. For simplicity, only one side of the absorption line is shown. Detailed description is in Section II-B.

average, the lower and upper mean values are averaged. ARTS performs simulations for the monochromatic pencil-beam intensity (or radiances). The simulated monochromatic intensities are converted to brightness temperatures according to the inverse Planck function. In Figs. 2 and 3, for simplicity, we show just one side of the WV-line. The values for the satellite T_B are calculated for both upper and lower passbands.

III. Results and Discussion

Fig. 2 shows the results of the closure study for the five radiometers and the five radiosonde stations. The matchup results for different MHS radiometers, installed on different satellite platforms (NOAA-19, MetOp-A, and MetOp-B), are presented separately, because there are known to exist small intersatellite biases [25]. The color of the markers indicates the radiosonde station and the marker type indicates the radiometer. For example, the "green star" at 184 GHz shows the weighted mean difference between the satellite and the radiosonde-simulated T_B for the 369 collocations that occurred from 2012 to 2015 for the radiosonde station Lindenberg and radiometer ATMS. The T_B difference is obtained using the weighted mean, as described in Section II. The red and green dashed lines present the mean value for all radiosounding stations and the MHS and ATMS radiometers, respectively. The errorbars show the 95% confidence interval based on the number of matchups.

As seen in Fig. 2, we find no significant closure gap: the radiosonde and the satellite measurements of WV agree within the measurement uncertainty limits. These limits are discussed later in this section. We find a systematic cold bias, the mean value for all the radiometers, of 0.42 K, but it is not enough to be significant.

The gray dashed line in Fig. 3 shows the results for the matchups between the radiosonde-simulated and the SAPHIR T_B from [3] ([4] obtained similar results in his study). The brown solid line is the corresponding averaged T_B difference obtained according to our methodology. We concentrate on this instrument separately, because it has the most channels around the WV-line and the spectral shape of the difference is the clearest. As seen in Fig. 3, our results differ significantly, especially for the outer channels. In order to identify the factors contributing to this difference, we carried out a sensitivity study. It takes into account a "reference" calculation, which includes all our processing steps, and "altered" calculations, which are executed by altering one parameter at a time. The parameters investigated are the radiosonde data processing type, the cloud filtering, and the methodology of comparison.

First, we investigate the influence of the improved GRUAN radiosonde data processing technique, we use for this paper, compared with the Vaisala data processing technique used in [3] and [4]. We use the approximations of the Vaisala-processed radiosonde profiles, according to [26, Fig. 17]. This figure shows that the difference between Vaisala and GRUAN processing increases with altitude. The data used to compute the effect are for the radiosonde-satellite pair: Manus and SAPHIR. Thus, the average effect of the radiosonde data processing type for the innermost channel is 1.0 K and it

goes down to 0.6 K for the outermost channel (blue arrows in Fig. 3).

Second, we look into the method of cloud detection in either of the measurements (the radiosonde or the satellite). As described in Section II, the previous studies used techniques based on the radiosonde relative humidity data and T_B differences between the channels (a method developed in [27]). Here, we use IR data from the SEVIRI instrument onboard MSG satellite. Because of the position of the MSG satellite, we can assess the impact of the IR cloud filter only for the Lindenberg station and the MHS and ATMS radiometers. MHS and ATMS radiometers have less channels than SAPHIR, that is why in Fig. 3, there is no estimate of the IR cloud filter effect for SAPHIR channels 1, 3, and 6. The average effect of the change in cloud filter is 0.04 K for the innermost channel and 0.40 K for the outermost channel (green arrows in Fig. 3). As the first, this second effect goes in the right direction for explaining the difference between this paper and the earlier ones.

Third, we analyze the impact of the different weightings for individual matchups, as part of the methodology of the comparison. The measure for the weighting is the inverse of the standard deviation squared of the satellite T_B inside the "target area." The effect of the weighting is 0.20 K for the innermost channel and 1.36 K for the outermost channel (orange arrows in Fig. 3). The data used to compute the effect is for two radiosonde-satellite pairs: Manus, Nauru, and SAPHIR. Even this effect goes in the same direction. In fact, for the outermost channels, it alone explains the difference between the different studies. If all the three considered effects would add up linearly, they would actually somewhat overcompensate the observed study differences. However, the radiative transfer is nonlinear, and moreover, some of the considered effects are not fully independent: for example, inhomogeneity (effect 3) is strongly correlated with cloudiness (effect 2). The overall conclusion from this sensitivity study is that the difference of our results to the previous studies can be well explained by the improvements in methodology.

Let us now return to the remaining closure gap between the radiosonde and the satellites. According to our analysis (Fig. 2), the mean value of the gap is 0.42 K. Although not statistically significant, as we will show later in the paragraph, the satellites do appear to be systematically colder in T_B and therefore moister than the radiosonde-based RTM simulations. We investigate the error budget for the satellite radiosonde difference to determine the effect of the main sources of error. The error budget has four parts: the radiosonde, the RT model, the satellite, and the collocation error. The results are shown in Table II and discussed in the following.

To estimate the real bias of the radiosonde, one has to compare its measurements with a more precise measurement system. Here, we use the frost point hygrometers, as in [26]. This paper found the mean error of the radiosonding data to be 10% of the WV volume mixing ratio for temperatures below 240 K. Taking this as an estimate of the uncertainty in the radiosondes due to systematic effects, we can map it to an uncertainty of 0.40 K for the innermost channel and of 0.10 K for the outermost channel in the error budget.

TABLE II
SYSTEMATIC ERROR BUDGET FOR THE SATELLITE RADIOSONDE DIFFERENCE

	Uncertainty range
Radiosonde	0.10–0.40 K
RT model	0.35–0.60 K
Satellite	0.38–0.77 K
Overall[1]	0.52–1.06 K

[1] The overall error budget was calculated as the square root of a sum of squares of all the parts of the uncertainties of the radiosonde, RT model and satellite.

The core of the RT model is the calculation of the absorption coefficient of the gasses. Precise spectroscopic data are needed for this calculation, among which the most important ones are the line strength, the "air" or "foreign" broadening and the continuum parameters. The spectroscopic parameters that we choose are very close to the ones recommended in [28] in a recent overview article, with a T_B difference amounting to only 0.01–0.05 K. Although the agreement is good, there is an uncertainty associated with the above-mentioned parameters. To estimate the effect of uncertainty of the spectroscopic data, we perturbed them with our estimate of their uncertainty: $+2\%$ for the line intensity, $+5\%$ for the "foreign" or "air" broadening, and $+25\%$ for the continuum. Such perturbations results in the error amounting to: 0.19–0.17 K for the line intensity, 0.01–0.30 K for the continuum for the innermost and the outermost channels, respectively. Also the uncertainty in the oxygen models (ARTS standard setup and [29]) influences the simulations for the WV-line, with an effect of 0.29–0.45 K for the innermost and the outermost channels, respectively. The estimated effect of the spectroscopy on the agreement, calculated as the square root of the sum of squares of individual components, is 0.35–0.60 K).

Preflight-tests and in-orbit performance evaluation give an estimation of the accuracy of the satellite data. The radiometric accuracy for the radiometers used in this paper is as follows: for ATMS 0.34–0.68 K [30], for SAPHIR 0.54–1.44 K [3], for MHS-NOAA-18 0.4–0.6 K, for MHS-MetOp-A 0.4–0.75 K, and for MHS-MetOp-B 0.25–0.35 K (Imke Hans, personal communication). The numbers are given for the outermost and innermost WV-channels, respectively. Another source of uncertainty is the spectral response of the radiometers and its representation in the RTM. The exact shape of the spectral response is more important when the spectrum changes, like in the vicinity of the absorption lines, and more so when the shape of the absorption line is not ideally symmetric and the upper and lower passbands sense slightly different values. The default parameter for the spectral response in RTM is rectangular: all the frequencies within the passband have the same weight during averaging. We conducted a sensitivity study to check the influence of the nonrectangular spectral response. The effect is 0.05–0.15 K for the outermost and innermost WV-channels, respectively (not shown in this paper). Based on the MHS sensitivity study, we can conclude that this effect is very small. On average, the effect of the satellite data on the agreement is 0.38–0.77 K.

The last term in the budget, the systematic error due to the imperfect collocation of the radiosonde and the satellite and the different spatial sampling, is more difficult to assess than the first three. There clearly is a large random error associated with this; in fact, this dominates the error for individual collocations, but we here assume that these errors average out and that the residual bias is zero.

The total combined effects of the three terms quantified previously are calculated as the square root of the sum of squares of individual components (Table II). All the above-mentioned factors amount to a bias of 0.52–1.06 K. This confirms the robustness of our conclusion: the satellite and the radiosonde measurements agree within the uncertainty limits.

However, one closure puzzle in the WV-line area remains: as discussed in [5, Fig. 2], assimilation residuals (so-called "observed minus simulated" difference) from numerical weather prediction models at the European Center for Medium-Range Weather Forecasts and Météo France also show a cold bias of the observations far from the line center of approximately 1–2 K. We have not investigated this issue here, but it would certainly be worthwhile to do so.

IV. Conclusion

In contrast to [3], [4], and [5], we find no significant closure gap between the satellite measurements at the WV-line and the radiosonde measurements. According to our analysis, the mean value of the gap is 0.42 K. The satellite measurements do appear to be systematically colder T_B-wise and therefore moister than the radiosonde-based RTM simulations, but the two agree within the measurement and methodology uncertainty (0.52–1.06 K). The sensitivity study suggests that the gap found in [3] and [4] is most likely caused by a combination of the radiosonde bias, residual cloud contamination, and atmospheric inhomogeneity.

Acknowledgment

The authors would like to thank the German Weather Service (DWD) for providing GRUAN-processed radiosounding data. They would also like to thank the NOAA CLASS Archive for providing the MHS and the ATMS data (http://www.class.ngdc.noaa.gov/). They would also like to thank the CNES/ISRO for providing the Megha-Tropiques Data. They would also like to thank the ICARE Data and Services Center for producing and distribution of Products and Images of SAPHIR data (http://www.icare.univ-lille1.fr). They would also like to thank EUMETSAT's Satellite Application Facility on Climate Monitoring for providing SEVIRI cloud mask data (DOI: 10.5676/EUM_SAF_CM/CMA_SEVIRI/V001). They would also like to thank O. Lemke for his assistance in getting the data. They would also like to gratefully acknowledge financial support through the Cluster of Excellence "CliSAP"(EXC177), Universität Hamburg, funded through the German Science Foundation (DFG). We thank Chris Merchant and other members of the EU H2020 project Fidelity and uncertainty in climate data records from Earth Observations for inspiring discussions on metrological aspects of satellite measurements. They would also like to thank the ARTS community for the longterm effort of developing this reference radiative transfer model, a crucial tool for closure studies like this one.

References

[1] V. Falcone *et al.*, "SSM/T-2 calibration and validation data analysis," Phillips Lab., Hanscom AFB, MA, USA, Tech. Rep. 1111, 1992, pp. 1–432.

[2] J. Li *et al.*, "International ATOVS processing package: Algorithm design and its preliminary performance," *Proc. SPIE*, vol. 3501, pp. 196–206, Aug. 1998.

[3] G. Clain, H. Brogniez, V. H. Payne, V. O. John, and M. Luo, "An assessment of SAPHIR calibration using quality tropical soundings," *J. Atmos. Ocean. Technol.*, vol. 32, pp. 61–78, Jan. 2015.

[4] I. Moradi, R. R. Ferraro, P. Eriksson, and F. Weng, "Intercalibration and validation of observations from ATMS and SAPHIR microwave sounders," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 11, pp. 5915–5925, Nov. 2015.

[5] H. Brogniez *et al.*, "A review of sources of systematic errors and uncertainties in observations and simulations at 183 GHz," *Atmos. Meas. Techn.*, vol. 9, no. 5, pp. 2207–2221, 2016.

[6] S. A. Buehler, M. Kuvatov, V. O. John, U. Leiterer, and H. Dier, "Comparison of microwave satellite humidity data and radiosonde profiles: A case study," *J. Geophys. Res. Atmos.*, vol. 109, no. D13, pp. 1–12, 2004.

[7] V. O. John and S. A. Buehler, "The impact of ozone lines on AMSU-B radiances," *Geophys. Res. Lett.*, vol. 31, no. 21, pp. 1–4, 2004.

[8] G. P. Anderson, S. A. Clough, F. X. Kneizys, J. H. Chetwynd, and E. P. Shettle, "AFGL Atmospheric Constituent Profiles (0.120 km)," Amer. FootGolf League, Palm Desert, CA, USA, Tech. Rep. TR-86-0110, 1986.

[9] S. Finkensieper *et al.* (2016). *CLAAS-2: CM SAF Cloud Property Dataset Using SEVIRI—Edition 2, Satellite Application Facility on Climate Monitoring*. [Online]. Available: https://doi.org/10.5676/EUM_SAF_CM/CLAAS/V002, doi: 10.5676/EUM_SAF_CM/CLAAS/V002

[10] G. Sèze, J. Pelon, M. Derrien, H. Le Gléau, and B. Six, "Evaluation against CALIPSO lidar observations of the multi-geostationary cloud cover and type dataset assembled in the framework of the Megha-Tropiques mission," *Quart. J. Roy. Meteorol. Soc.*, vol. 141, no. 688, pp. 774–797, 2015.

[11] (2017). *GCOS Reference Upper-Air Network*. [Online]. Available: https://www.gruan.org/

[12] K. Dieter Klaes and R. Schraidt, "The European ATOVS and AVHRR processing package (AAPP) development," in *Proc. Int. ATOVS Study Conf. (ITSC-10)*, 1999, pp. 288–294.

[13] J. Robel *et al.*, "NOAA KLM user's guide with NOAA-N, -N' supplement," NOAA NESDIS NCDC, Asheville, NC, USA, Tech. Rep., 2009.

[14] I. Hans, M. Burgdorf, V. O. John, J. Mittaz, and S. A. Buehler, "Noise performance of microwave humidity sounders over their lifetime," *Atmos. Meas. Tech.*, vol. 10, pp. 4927–4945, Dec. 2017.

[15] H. Brogniez, P.-E. Kirstetter, and L. Eymard, "Expected improvements in the atmospheric humidity profile retrieval using the Megha-Tropiques microwave payload," *Quart. J. Roy. Meteorol. Soc.*, vol. 139, no. 673, pp. 842–851, 2013.

[16] R. Roca *et al.*, "The Megha-Tropiques mission: A review after three years in orbit," *Frontier Earth Sci.*, vol. 3, p. 17, May 2015.

[17] NASA. (2017). *Advanced Technology Microwave Sounder (ATMS)*. [Online]. Available: https://jointmission.gsfc.nasa.gov/atms.html

[18] S. A. Buehler, P. Eriksson, T. Kuhn, A. von Engeln, and C. Verdes, "ARTS, the atmospheric radiative transfer simulator," *J. Quant. Spectroscopy Radiat. Transf.*, vol. 91, no. 1, pp. 65–93, 2005.

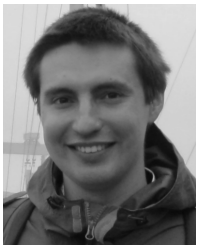
[19] P. Eriksson, S. A. Buehler, C. P. Davis, C. Emde, and O. Lemke, "ARTS, the atmospheric radiative transfer simulator, version 2," *J. Quant. Spectroscopy Radiat. Transf.*, vol. 112, no. 10, pp. 1551–1558, 2011.

[20] M. Kuntz, "A new implementation of the Humlicek algorithm for the calculation of the Voigt profile function," *J. Quant. Spectroscopy Radiat. Transf.*, vol. 57, no. 6, pp. 819–824, 1997.

[21] J. H. van Vleck and D. L. Huber, "Absorption, emission, and linebreadths: A semihistorical perspective," *Rev. Modern Phys.*, vol. 49, no. 4, pp. 939–959, 1977.

[22] V. H. Payne *et al.*, "Air-broadened half-widths of the 22- and 183-GHz water-vapor lines," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 11, pp. 3601–3617, Nov. 2008.

[23] E. J. Mlawer, V. H. Payne, J.-L. Moncet, J. S. Delamere, M. J. Alvarado1, and D. C. Tobin, "Development and recent evaluation of the MT_CKD model of continuum absorption," *Philos. Trans. Roy. Soc. A*, vol. 370, no. 1968, pp. 2520–2556, 2012.
[24] P. Eriksson, M. Ekström, C. Melsheimer, and S. A. Buehler, "Efficient forward modelling by matrix representation of sensor responses," *Int. J. Remote Sens.*, vol. 27, no. 9, pp. 1793–1808, 2006.
[25] V. O. John, G. Holl, S. A. Buehler, B. Candy, R. W. Saunders, and D. E. Parker, "Understanding intersatellite biases of microwave humidity sounders using global simultaneous nadir overpasses," *J. Geophys. Res.*, vol. 117, no. D2, pp. 1–13, 2012.
[26] R. J. Dirksen, M. Sommer, F. J. Immler, D. Hurst, R. Kivi, and H. Vömel, "Reference quality upper-air measurements: GRUAN data processing for the Vaisala RS92 radiosonde," *Atmos. Meas. Tech.*, vol. 7, no. 12, pp. 4463–4490, 2014.
[27] S. A. Buehler *et al.*, "A cloud filtering method for microwave upper tropospheric humidity measurements," *Atmos. Chem. Phys.*, vol. 7, no. 21, pp. 5531–5542, 2007.
[28] M. Y. Tretyakov, "Spectroscopy underlying microwave remote sensing of atmospheric water vapor," *J. Molecular Spectroscopy*, vol. 328, pp. 7–26, Oct. 2016.
[29] M. Y. Tretyakov, M. A. Koshelev, V. V. Dorovskikh, D. S. Makarov, and P. W. Rosenkranz, "60-GHz oxygen band: Precise broadening and central frequencies of fine-structure lines, absolute absorption profile at atmospheric pressure, and revision of mixing coefficients," *J. Molecular Structroscopy*, vol. 231, pp. 1–14, May 2005.
[30] E. Kim, C.-H. J. Lyu, K. Anderson, R. V. Leslie, and W. J. Blackwell, "S-NPP ATMS instrument prelaunch and on-orbit performance evaluation," *J. Geophys. Res., Atmos.*, vol. 119, no. 9, pp. 5653–5670, 2014.



Oleksandr Bobryshev received the M.S. degree in meteorology from the Taras Shevchenko National University of Kyiv, Kiev, Ukraine, in 2014. He is currently pursuing the Dr.rer.nat. degree with the Meteorological Institute, Universität Hamburg, Hamburg, Germany. His M.S. thesis was entitled Retrieval Profiles of Temperature and Humidity Based on the Satellite Data NOAA/ATOVS.

From 2011 to 2014, he was a first class Engineer with the Ukrainian Hydrometeorological Institute, Kyiv. His research interests include radiative transfer simulations for operational microwave satellite sensors and the evaluation of models using observations.

Stefan A. Buehler received the M.S. degree in physics from The State University of New York, Stony Brook, NY, USA, in 1994, and the Ph.D. degree in physics from the University of Bremen, Bremen, Germany, in 1998.

From 2006 to 2013, he was a Professor Chair with the Luleå University of Technology, Kiruna, Sweden. He is currently a Professor of applied meteorology and the Director of the Meteorological Institute, Universität Hamburg, Hamburg, Germany. His research interests include radiative transfer modeling for remote sensing and radiative energy fluxes, numerical methods, new sensor concepts for water vapor and cloud ice remote sensing, satellite climate data records, cloud ice microphysics, and the role of water vapor and clouds in the climate system.

Viju O. John received the M.S. degree in physics and the M.S. degree in atmospheric science from the Cochin University of Science and Technology, Kochi, India, in 1998 and 2000, respectively, and the Ph.D. degree in physics from the University of Bremen, Bremen, Germany, in 2005.

He is currently an fidelity and uncertainty in climate data records from Earth observations Scientific Expert with the European Organisation for the Exploitation of Meteorological Satellites, Darmstadt, Germany, and a Senior Scientist with the Met Office Hadley Centre, Exeter, U.K. His research interests include the remote sensing of water vapor in the Earth's atmosphere, intercalibration of satellite sensors, evaluation climate models using satellite observations, and generation of climate data records of essential climate variables tailored for climate monitoring and for assimilation in climate reanalyses.

Manfred Brath received the Diploma degree in physics from the Universität Hamburg, Hamburg, Germany, in 2008, and the Ph.D. degree from Universität Hamburg in 2013, with a focus on the active microwave remote sensing of sea ice. His Diploma thesis was on eddy kinetic energy and eddy transports using Jason-TOPEX/Poseidon tandem mission data.

He is currently a Senior Scientist with the Meteorological Institute, Universität Hamburg. His research interests include microwave remote sensing of clouds, microwave radiative transfer, and electromagnetic scattering of hydrometeors.

Hélène Brogniez received the M.S. degree (DEA) and the Ph.D. degree in physics of remote sensing from Paris 6 University, Paris, France, in 2001 and 2004, respectively.

She is currently an Associate Professor with the University of Versailles Saint Quentin-en-Yvelines, Versailles, France. Her research interests include the passive remote sensing of atmospheric water vapor, tropical, and subtropical climates, and model answers to climate change.