Earth and Space Science

TECHNICAL REPORTS: METHODS

10.1029/2019EA000856

Opportunistic Constant Target Matching—A New Method for Satellite Intercalibration

Stefan A. Buehler[1], Marc Prange[1,4], John Mrziglod[2], Viju O. John[3], Martin Burgdorf[1], and Oliver Lemke[1]

[1]Meteorologisches Institut, Centrum für Erdsystem- und Nachhaltigkeitsforschung (CEN), Fachbereich Geowissenschaften, Universität Hamburg, Hamburg, Germany, [2]Φ-Lab, European Space Agency (ESA) Hamburg, Germany, [3]European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Frascati, Italy, [4]International Max Planck Research School on Earth System Modelling (IMPRS-ESM) Darmstadt, Germany

Key Points:

- Opportunistic constant target matching is a new method for satellite cross-calibration
- We confirm the almost negligible bias between NOAA 18 and Metop A MHS Channel 3 (below 0.05 K)

Correspondence to:
S. A. Buehler,
stefan.buehler@uni-hamburg.de

Citation:
Buehler, S. A., Prange, M., Mrziglod, J., John, V. O., Burgdorf, M., & Lemke, O. (2020). Opportunistic constant target matching—A new method for satellite intercalibration. *Earth and Space Science. 7*, e2019EA000856. https://doi.org/10.1029/2019EA000856

Received 23 AUG 2019
Accepted 23 MAR 2020
Accepted article online 14 APR 2020

Abstract Opportunistic constant target matching is a new method for satellite intercalibration. It solves a long-standing issue with the traditional simultaneous nadir overpass (SNO) method, namely, that it typically provides only data points with cold brightness temperatures for humidity sounding instruments on sun-synchronous satellites. In the new method, a geostationary infrared sensor (SEVIRI) is used to select constant target matches for two different microwave sensors (MHS on NOAA 18 and Metop A). We discuss the main assumptions and limitations of the method and explore its statistical properties with a simple Monte Carlo simulation. The method was tested in a simple case study with real observations for this combination of satellites for MHS Channel 3 at 183 ± 1 GHz, the upper tropospheric humidity channel. For the studied 3-month test period, real observations are found to behave consistently with the simulations, increasing our confidence that the method can be a valuable tool for intercalibration efforts. For the selected case study, the new method confirms that the bias between NOAA 18 and Metop A MHS Channel 3 is very small, with absolute value below 0.05 K.

1. Introduction

Intercalibrating sensors on different satellites is one of the greatest challenges in observational climate science, since global and continuous data sets spanning decades can only be constructed from multiple satellites (e.g., Hans et al., 2019; Mears & Wentz, 2005; Ohring et al., 2005; Shi & Bates, 2011; Scott et al., 1999; Tabata et al., 2019; Zou & Wang, 2011, among many others). This article presents a new method to obtain intercalibration points for such undertakings. The most established method, simultaneous nadir overpasses (SNOs), has several shortcomings, most importantly that it yields only cold (in terms of brightness temperature) calibration points for many important observation frequencies, whereas an optimal intercalibration requires also warm calibration points, so that the full dynamical range of the involved sensors is covered.

The new method does not require the two sensors to observe the same target at the same time. Instead, we opportunistically identify cases where the atmosphere has been stable between matches that can be several hours to days apart in time. Geostationary satellite imagery is used for identifying these cases. As shown below, the new method has the potential to allow for the first time a precise intercalibration of any simultaneously operational sounding satellites, a problem for which hitherto no method without severe shortcomings existed.

So far, the gold standard for intercalibrating sensors on different satellites was SNO matchups (Cao et al., 2004, 2005; Holl et al., 2010; Iacovazzi & Cao, 2007; John et al., 2012; John, Allan, et al., 2013; John, Holl, et al., 2013; Nagle & Holz, 2009; Shi et al., 2008; Wang et al., 2007; Zou et al., 2006). As discussed in detail for example in John et al. (2012), these occur only in high latitudes when matching two Sun-synchronous satellites, except for very rare cases where satellites have similar equator crossing times for a short time period due to orbit drift. At these high latitudes, humidity sounding channels see the surface, which they normally do not. Especially for microwave instruments, this makes the brightness temperatures (T_Bs) very cold, since microwave emissivity is well below 1. As a consequence, SNO intercalibration matchups for microwave humidity channels are all at the cold end of the overall observed brightness temperature range (see Figure 1). This is a severe problem for sensor harmonization efforts, such as the ones that have been

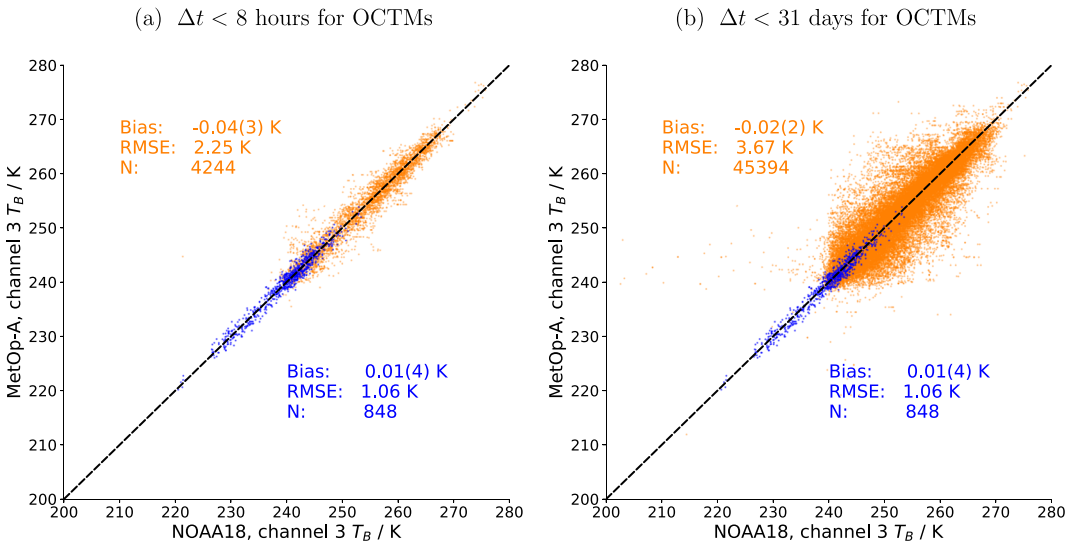

Figure 1. Matchup scatterplot of NOAA 18 versus Metop A for two different temporal collocation criteria applied on OCTMs. Orange points in both plots mark OCTMs. Blue points are the same in (a) and (b) and mark SNO matchups obtained when applying the established collocation criteria of $\Delta t < 5$ min and $\Delta s < 5$ km (John et al., 2012). Orange and blue text gives similar statistics as in Figure 5 (total number of cases, standard deviation, and the bias with its standard uncertainty in parentheses).

made within the EU-funded project FIDUCEO (Fidelity and uncertainty in climate data records from Earth Observations, https://www.fiduceo.eu).

The new idea that is explored in this work is to relax the condition that both sensors must observe at the same time and replace it by the condition that the atmosphere must not have changed between the two observations. As shown below, it is indeed possible to find such cases, even many hours or days apart, using imagery from geostationary sensors.

We test the new method by intercomparing Channel 3 of the Microwave Humidity Sounder (MHS) on the satellites NOAA 18 and Metop A, using as reference Channel 5 of the geostationary Spinning Enhanced Visible and Infrared Imager (SEVIRI) on the Meteosat Second Generation (MSG) satellite. This particular microwave sounder pair was chosen because a previous study (Hans et al., 2019) has shown them to be among the best calibrated and most stable microwave sensors, arguably agreeing within the accuracy limitations of the comparison technique. Focusing on a sensor pair with no known bias provides for a stronger test of the new method than application to a sensor pair with a large bias would.

The structure of the article is as follows: To set the scene, section 2 discusses some statistical considerations. Section 3 presents an example that illustrates the properties of these quasi-stable targets. Section 4 then describes the adopted matchup procedure, including the chosen filter thresholds. Section 5 presents a Monte Carlo test of the new method, followed by a case study with real data in section 6. Implications of the new method are discussed in section 7, and then the paper ends with summary and conclusions in section 8.

2. Statistical Considerations

Before delving into the new matchup method, some statistical considerations are in order. For simplicity, we here assume that the statistics between the two instruments to be compared can be modeled by a stationary stochastic process, in other words that their mean difference and its standard deviation do not change over time. This assumption is often not valid for real instruments, an issue that we will return to at the end of this section.

Under this assumption of stationarity, the precision σ_M, to which one can determine the mean difference between two instruments from a set of uncorrelated measurements follows from the law of error propagation to be

$$\sigma_M = \frac{\sigma}{\sqrt{N}}, \tag{1}$$

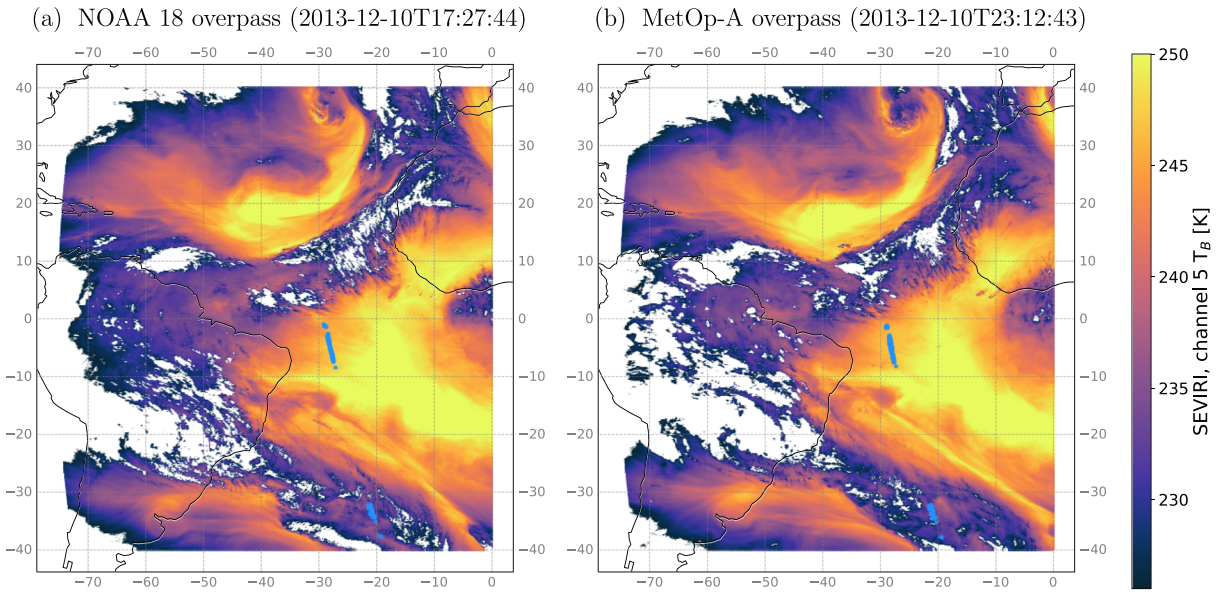

Figure 2. An example for opportunistic constant target matchups. (a) The geostationary SEVIRI image at 6.2, μm at the time of the NOAA 18 overpass; (b) the same area approximately 6 hr later at the time of the Metop A overpass. Blue markers in both plots show the identified nadir matchup pixels of the respective satellites. White areas in both plots mark regions with high clouds, where the brightness temperature is below the color bar range.

where σ is the standard deviation of the difference and N is the number of measurements. This implies that σ_M improves only slowly with N for large N and that one therefore needs a perhaps somewhat surprisingly large number of measurements to accurately determine a bias that is small relative to the variability σ. The number is given by

$$N = \left(\frac{\sigma}{\sigma_M}\right)^2. \tag{2}$$

The bias that one is hoping to correct in satellite harmonization is of the order of 0.1 K, so a reasonable target value for σ_M is 0.01 K. If one just averages microwave data over the entire tropics, the σ due to natural variability is approximately 8 K (for the innermost 183 GHz channel), implying a σ for intersatellite differences of $8\sqrt{2}$ K = 11.3 K. According to equation (2), this means that simple random matching requires a bit more than one million individual datapoints to reach acceptable bias precision.

If observations are matched up, as in the traditional SNO method, the variability of the intersatellite difference is substantially lower. With a typical σ of 1 K, only 10,000 individual datapoints are needed to reach the same 0.01 K bias precision. But it may still be a challenge to find so many SNOs, in particular if mission overlap time periods are short.

These relatively large numbers may be the reason for some of the observed discrepancies between intercalibration attempts, since with fewer observations the determined bias simply has a significant random error. Viewed only from the point of view of sampling statistics, simple areal averaging methods could be superior to matched up data sets, because their advantage of a much higher number of data points outweighs their disadvantage of a higher σ. However, their biggest problem is that sampling effects can easily alias into a bias. Notably, if the brightness temperature on average has a diurnal cycle, it will alias into an intersatellite bias, if different satellites measure at different local time. It will be important to check how our new method behaves in this respect.

Finally, we have to remember that biases between real instruments may not be constant in time. In that case the intercalibration has to be done for suitable time windows, and the shorter these windows the fewer the number of datapoints that can be used for intercalibration. This can in practice further aggravate the problem of poor intercalibration statistics.

3. Opportunistic Constant Targets

Since SNOs occur typically at high latitudes, additional matchups are most urgently needed for low latitudes. The tropical atmosphere is characterized by relatively small convective areas that change rapidly, and large

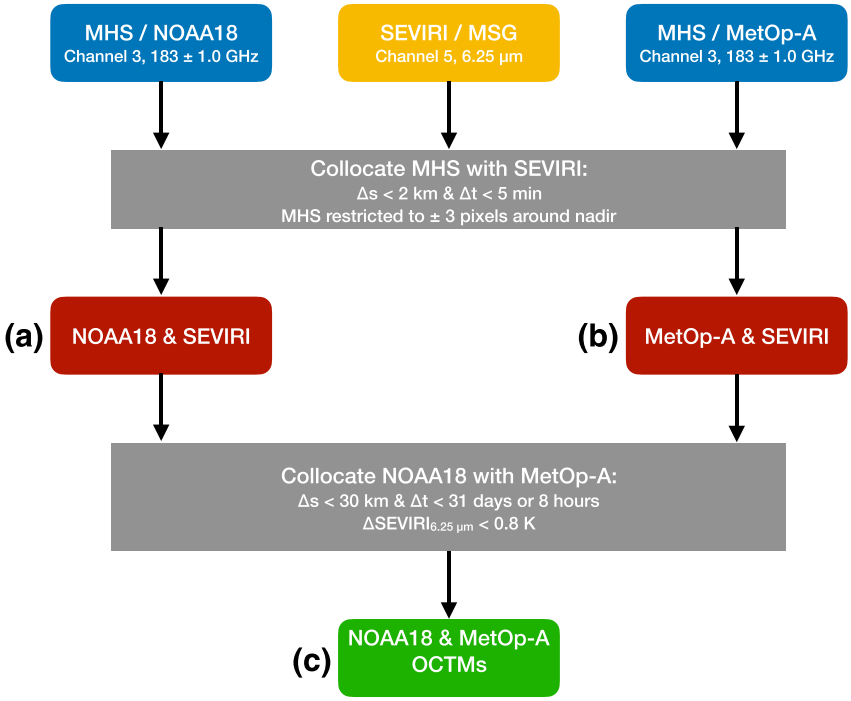

Figure 3. Algorithm flowchart for finding OCTMs; see text for details.

subsiding areas that change and move slowly. Figure 2 illustrates this. It shows two images from Channel 5 at 6.2 µm of the Spinning Enhanced Visible and Infrared Imager (SEVIRI) instrument on the geostationary Meteosat Second Generation (MSG; Schmetz et al., 2002) satellite. The large subsidence zone, characterized by orange colors, has moved only slightly in the approximately 6 hr between the two images.

These two images were selected because they are coincident with the overpasses of two polar orbiting meteorological satellites, NOAA 18 for Image (a) and Metop A for Image (b). The idea of the new method is to use the SEVIRI images to identify pixel pairs for the microwave sounders on the polar orbiting satellites for which the atmosphere is relatively invariant. We call these pixel pairs "Opportunistic Constant Target Matchups," in short OCTMs. The locations of the 93 OCTMs for these two scenes are also marked in Figure 2. As the figure hopefully helps to convey, OCTMs occur quite naturally (and frequently). The exact collocation algorithm is described in the next section. Implications and limitations of this new approach are also discussed further down.

4. Matchup Procedure

The actual matchup was done with a general collocation software package developed by one of the authors (John Mrziglod), and freely available as part of the "Typhon" collection of Python functions for atmospheric science (https://www.radiativetransfer.org/tools/). The collocation algorithm, described in detail in Mrziglod (2018), uses a ball tree (Dolatshah et al., 2015) as a spatial index for one of the two data sets, in order to make the matching efficient, as suggested by Bentley and Maurer (1980). It is loosely based on the collocation algorithm by Holl et al. (2010).

Figure 3 illustrates the complete opportunistic constant target matchup algorithm for the case of the Microwave Humidity Sounder (MHS) on NOAA 18 and on Metop A. For simplicity, we consider only Channel 3, at 183 ± 1 GHz, an upper tropospheric humidity sounding channel, with vertical sensitivity broadly in the 500–200 hPa region.

First, separate matchup data sets of the two MHS instruments against SEVIRI on MSG are created. So far, we consider only the six innermost pixels of the MHS scan, with viewing angle close to nadir. The spatiotemporal matchup criteria with SEVIRI are $\Delta s < 2$ km and $\Delta t < 5$ min. This results in Data Set A, which contains matching pairs of SEVIRI and NOAA 18 observations, and Data Set B, which contains matching pairs of SEVIRI and Metop A observations.

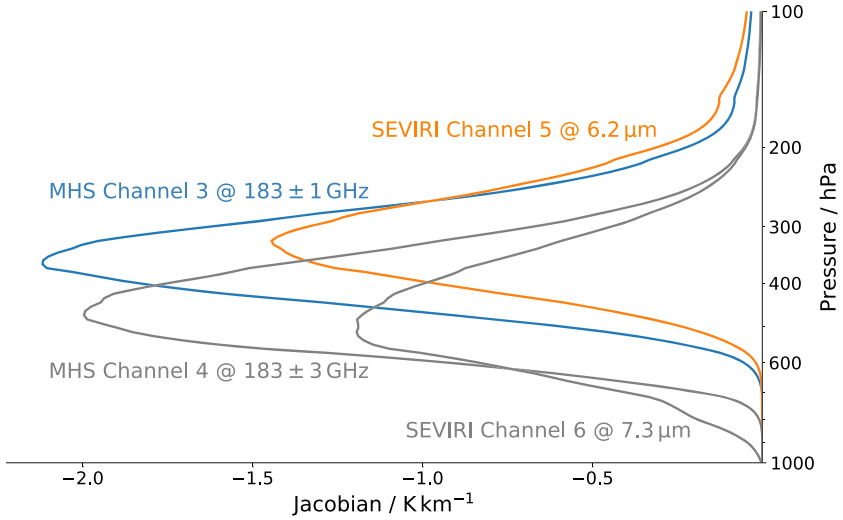

Figure 4. Water vapor Jacobians for MHS Channel 3 (183 ± 1 GHz, blue) and SEVIRI Channel 5 (6.25µm, orange) for a tropical standard atmosphere (FASCOD; Anderson et al., 1995) and the nadir viewing direction. The water vapor perturbation is in fractional units (∗ 1.1 = 10% increase) and is normalized by the geometrical layer thickness in kilometers, so that numerical values roughly represent the change in brightness temperature for a doubling of the water vapor concentration in a 1 km thick layer. Calculations were performed with the Atmospheric Radiative Transfer Simulator (ARTS; Buehler et al., 2018), using the predefined setups for MHS and SEVIRI. For comparison, Jacobians for MHS Channel 4 (183 ± 3 GHz) and SEVIRI Channel 6 (7.3 µm) are also shown in gray.

In a second step, these two data sets are matched against each other with only a loose spatial and temporal matchup constraint ($\Delta s < 30$ km and $\Delta t < 8$ hr or even much longer as discussed below), but under the additional constraint that the difference in SEVIRI Channel 5 (6.25µm) brightness temperature of Data Sets A and B is less than 0.8 K. The rationale for this threshold number is discussed in section 5. The result is Data Set C, which contains matching quartets of two SEVIRI observations (which differ from each other by less than 0.8 K), one NOAA 18 observation, and one MHS observation. We then proceed to compare the NOAA 18 and MHS observations from these quartets.

The SEVIRI observations are not used further but have served their purpose as "filter observations" to select matching observations from the microwave instrument pairs. SEVIRI Channel 5 was selected, because it also is a humidity sounding channel, roughly sensitive in the same altitude range as MHS Channel 3. This is illustrated by Figure 4, showing water vapor Jacobians for MHS Channel 3 (183 ± 1 GHz) and SEVIRI Channel 5 (6.25µm). Although the Jacobians are quite similar, it is important to note that there are important physical differences between the MHS and SEVIRI observations. Perhaps most importantly, MHS observes the flanks of a single strong water vapor spectral line, whereas SEVIRI observes a band that includes many such individual lines, with regions of weaker absorption in-between. In addition to this, the viewing angle of each instrument (for the individual pixels compared) also matters, as the Jacobian peak moves to higher altitude for more oblique viewing angle. These differences prevent us from doing traditional double differencing between SEVIRI and MHS.

Instead, we make only the weaker key assumption that if the atmosphere is unchanged according to the SEVIRI channel, it will also be unchanged, or very nearly so, for the MHS channel that is to be matched up. Note that this assumption is considerably less strict than the traditional double-differencing assumption, which would be that the difference of MHS to SEVIRI is a proxy for the difference to another MHS instrument. Our weaker assumption will hold also if SEVIRI and MHS have a different nonlinear mapping to the underlying atmospheric state (as long as they are sensitive to roughly the same altitude region).

The overall result of this procedure is a set of matched NOAA 18/Metop A MHS Channel 3 brightness temperatures, for which we have good reason to assume that both instruments observe the same target, even though the time of observation may differ by several days. (The two matchup time cutoffs we have explicitly investigated are 8 hr and 31 days, as discussed in section 7.)

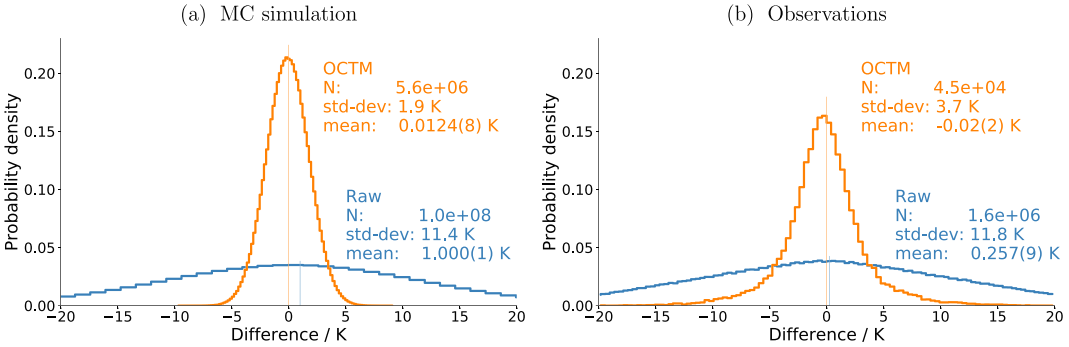

Figure 5. LEO-LEO difference before and after OCTM matching. The blue curve is the normalized difference histogram between random afternoon and morning measurements, showing a broad distribution due to natural variability and a mean difference due to the diurnal cycle. The orange curve is the LEO-LEO difference after (GEO-assisted) OCTM matching. It is based on fewer individual cases, since the pairs with GEO differences above the matching threshold are removed. Its variability (width) is much smaller, and its mean is 2 orders of magnitude smaller than for the unfiltered data. Thin vertical lines indicate the mean of each histogram. The number of cases in each histogram (N), their standard deviation, and their mean value are indicated next to the curves. Numbers in parentheses represent standard uncertainty expressed in terms of the least significant digit of the mean. (a and b) The results of the MC-simulation and of real observations, respectively.

Without loss of generality, the test region for this pilot study was somewhat arbitrarily restricted to −40° to 40° latitude and −70° to 0° longitude to make the data volume more easily manageable. In other words, we use only the western half of the full SEVIRI image, and only the low latitudes. This region is large enough to get a large number of collocations (about 45,400 for these two satellites) in the 3-month time period of the study from 1 October 2013 to 31 December 2013. Low latitudes were selected because additional collocations are most urgently needed there. All pixels in the test area were considered for matching, with no distinction between land and ocean or nighttime and daytime data.

5. Monte Carlo Test

A simple Monte Carlo (MC) simulation was set up, in order to better understand the statistical properties of the new method. In particular, we want to quantify to what extent a diurnal cycle in the raw brightness temperatures will alias into the matchup differences.

The MC simulation starts with two random sequences *true_morning* and *true_afternoon* that each contain Gaussian random numbers with $\sigma = 8$ K. The mean value of *true_morning* is arbitrarily chosen to be 300 K (this value does not affect the results), and the mean value of *true_afternoon* is 1 K higher, mimicking a diurnal cycle. The blue curve in Figure 5a shows a histogram of the *leo_afternoon-leo_morning* differences. By construction, they have a mean value of 1 K (the diurnal cycle bias) and a σ of $8\sqrt{2}$ K.

From each of the two *true* … random sequences, two other sequences are generated by adding additional noise: *leo_morning* is generated by adding Gaussian noise with $\sigma = 1$ K and zero mean to *true_morning*. It mimics the polar orbiting microwave sensor (LEO stands for low Earth orbit). Similarly, *geo_morning* is generated from *true_morning* by adding noise with $\sigma = 0.8$ K, mimicking the geostationary SEVIRI measurement. The … *afternoon* leo and geo sequences are generated in the same fashion. (Note that these noise assumptions are based on prelaunch specifications and thus are quite conservative, especially the SEVIRI instrument may have significantly lower noise in practice. We will come back to this in the next section.)

Now the matchup procedure is simulated by comparing the *geo_morning* and *geo_afternoon* sequences, and selecting cases where their absolute difference is below 0.8 K. We set the matchup window size to be the same as the SEVIRI noise, which is a good compromise that yields a high number of matches without broadening their difference distribution too much. A sensitivity study (not shown) showed that results do not depend strongly on the exact threshold number.

For the selected matches, we then examine the *leo_afternoon-leo_morning* differences. The results are displayed as the orange curve in Figure 5a. Their much smaller σ results from the combination of SEVIRI noise, microwave noise, and the width of the matchup window. Most notably, their mean (the remaining diurnal cycle bias) is only 0.01 K.

We conclude that the OCTM method has the potential to almost completely eliminate a diurnal cycle bias in the raw data. The large 8 K natural variability actually helps here, since it makes the raw difference distribution quite flat in the range from which the OCTM matches are drawn (the blue curve is flat in the center, where the orange curve sits).

The biggest limitation in this simple MC test is that the natural variability was assumed to be Gaussian. For other distributions the aliasing effect may be larger, but at least the test shows that the new method is statistically reasonable.

6. Real Observations

In the previous section, it was shown that OCTMs are capable of removing a predetermined diurnal cycle in the MC model. Now it is interesting to take a look at how the method performs when applied to real observations. Figure 5b depicts the probability density distributions of LEO differences based on Metop A and NOAA 18 data. As in the MC case, the blue curve is based on randomly combining all observations within the chosen test area and period. It can be viewed as a reference to the orange curve, which is based on OCTMs obtained from the same test period and area when applying the collocation criteria of $\Delta s < 30$ km, $\Delta t < 31$ days, $\Delta SEVIRI < 0.8$ K.

The validity of the premise that constant SEVIRI observations imply constant MHS observations can be assessed by comparing the OCTM distribution of the real observations to the MC simulation. In the simulated ideal case, where it is assumed that SEVIRI is perfectly representative of MHS, except for radiometric noise, the standard deviation of the matched data is 1.9 K. For the real observations it is slightly higher, 3.7 K, but far below the unmatched raw data (11.8 K). We attribute the higher value in the real observations to "representation noise" resulting from imperfect horizontal alignment, different footprint sizes, and different vertical sampling characteristics of the two instrument classes. If one were to assume that the total noise on the matched data results from a squared combination of the noise from the idealized simulation and this representation noise, and all these errors were Gaussian, it would imply a value of the representation noise of $\sqrt{3.7^2 - 1.9^2} = 3.2$ K. In reality, however, the representation noise is super-Gaussian on the wings, which means that comparatively rare outliers contribute quite strongly to the calculated standard deviation. Overall, we conclude that these statistics demonstrate that the premise of constant SEVIRI implying constant MHS holds reasonably well.

As mentioned in section 5, especially, our SEVIRI noise assumptions are very conservative. We therefore did an additional Monte Carlo simulation with tighter specifications of 0.5 K MHS noise and 0.05 K SEVIRI noise. These specifications are based on a recent conference article for MHS (Bonsignori, 2007) and on a public technical report from EUMETSAT for SEVIRI (EUMETSAT, 2019). Despite the lower instrument noise assumptions, the matchup window was kept at 0.8 K, so that it still matches our processing of the real observations. The lower noise simulation results in a standard deviation of the matched data of 0.8 K, lower than the 1.9 K that we obtain in our reference simulation with the prelaunch specifications. However, following the lines of the above discussion, this would imply an only slightly higher representation noise value of $\sqrt{3.7^2 - 0.8^2} = 3.6$ K, so the above conclusion still holds. In any case, with both sets of noise assumptions, the representation noise dominates over the instrument noise. This matches also our general experience with other instrument intercomparison techniques; see, for example, the discussion for the case of collocated total column water vapor measurements in Buehler et al. (2012).

Concerning the actual bias between the two MHS sensors, the main issue when assessing the results of real observations compared to the results of the MC model is that there is only poor quantitative knowledge about the true diurnal cycle and the true intersatellite bias. Hence, based on real observations, it is impossible to judge the absolute accuracy of the method, for example, to which degree it is capable of determining the pure instrument-related intersatellite bias. However, the obtained biases from real observations appear reasonable in the sense that a 1 order of magnitude larger bias exists in the randomly matched data set compared to the OCTM data set, which is most likely related to the atmospheric diurnal cycle. This is in qualitative agreement with previous studies based on other methods that also did not indicate strong intersatellite bias between Metop A and NOAA 18 (Hans et al., 2019; John et al., 2012).

The computed standard errors obtained from the real observations given in the brackets in Figure 5b reflect the precision of the method. The standard error of the OCTM data set is roughly doubled compared to the

randomly matched data set. The main reason for this is the 2 orders of magnitude lower data amount of OCTMs compared to the randomly matched data. However, the achieved precision of 0.02 K by the OCTMs is a sufficiently low value to reliably determine biases in the order of 0.1 K as desired for satellite harmonization purposes.

It can be concluded that the OCTM matchup procedure statistically behaves as expected from the simple MC simulation. While it is impossible to judge from the real observations alone whether the method reveals the correct intersatellite bias, the MC simulation showed that the method is capable of removing undesired diurnal cycle aliasing. The precision of the method in our case study easily suffices for satellite harmonization purposes with a value of 0.02 K. Note, however, that the precision can deviate when applying the method to other channels and instruments and that particularly the amount of OCTMs must be large enough to allow for a robust analysis.

7. Discussion

The new method basically is a special form of double differencing. By doing close collocations of the two microwave sensors with the infrared sensor SEVIRI, and then focusing on those measurements that stay constant over a longer time period, it minimizes the impact of sampling errors. The method makes only relatively weak assumptions.

The first assumption is that the chosen SEVIRI and microwave channels have sufficiently similar sounding altitude so that nearly constant SEVIRI observations imply nearly constant microwave observations. This assumption is weaker than the traditional assumption in double differencing, which one observation is representative of the other over the complete dynamical range. Traditional double differencing has been applied by Wang et al. (2010) to intercalibrate different infrared sounders using an infrared geostationary transfer instrument, but to our knowledge, it has never been attempted to intercalibrate microwave instruments with an infrared transfer instrument.

Another way in which double differencing has been applied to satellite intercalibration is by using forward-modeled observations from a numerical weather prediction model as the transfer observation (Chung, Soden, & John, 2013; Saunders et al., 2013). This is sometimes referred to as the O-B (observation minus background) method. The key assumption in that case is that the model does not have diurnal cycle biases, since these would map into apparent instrument biases when comparing satellites with different overpass local times.

While our new method does not rely on a model correctly representing the diurnal cycle, it does depend on the second obvious and important assumption that the SEVIRI sensor is stable for the several hour or even days time period between the microwave matchups. More specifically, it does not have any systematic difference between the two observation times, since random errors just contribute to the noise in the matchup data and thus are no problem. If SEVIRI had a systematic error that followed a diurnal cycle, it would be detrimental for the method. As discussed in Hewison et al. (2013), this should not be an issue for spin stabilized geostationary satellite instruments such as SEVIRI, but it could be an issue for some of the older sensors on three-axis stabilized platforms (Yu et al., 2013).

A third key assumption is that the diurnal cycle is small compared to the natural variability (which is fulfilled given the small amplitude of the diurnal cycle for sounding channels reported in other studies; Kottayil et al., 2013, 2016; Chung et al., 2007; Chung, Soden, Sohn, et al., 2013) and that in consequence the distribution of the variability inside the 0.8 K SEVIRI matchup thresholds is unbiased. We tested this by examining the SEVIRI brightness temperature difference between the two microwave matchup pair members (not shown), and it is indeed uniform between −0.8 and +0.8 K, with the exception of a modest spike at zero due to SEVIRI numeric truncation noise, which has no impact on the results.

The above assumptions seem to be the most critical ones for the new method. They seem to hold in the examined test case and thus lend some confidence in the methods suitability for the targeted application of harmonizing microwave sounder data. They will have to be carefully checked when adapting the method for other applications.

The big strength of the new method is that it can provide warm matchup points (collocated measurements where the observed brightness temperature is high), solving a critical issue for intercalibration exercises.

Figure 1 illustrates this by a scatter plot of the new OCTM matches together with the traditional SNO matches, for the same case study discussed above. With SNOs alone it is not possible to properly intercalibrate the two sensors, because their bias at warm brightness temperatures in unconstrained. With both techniques combined, one can achieve good coverage of the entire dynamical range. The noise level of OCTMs is two to three times that of SNOs, but this is still acceptable for practical use.

The figure also illustrates the impact of the time matchup threshold between the two microwave instruments. The two cases shown are (a) 8 hr and (b) 31 days. The extreme 31-day case was included because it generates significantly more matches. The variability (RMSE of intersatellite difference) with this very large window is modestly higher (3.67 K instead of 2.25 K for the 8 hr window), but the bias is smaller and the bias precision similar (−0.02(2) K instead of −0.04(3) K). Although the difference in bias is not statistically significant in this case, the advantage of more matches may outweigh the disadvantage of higher noise in some cases. Judging from the very small resulting bias, SEVIRI appears to be sufficiently stable to allow this very long matchup time window. Future applications will have to do their own trade-off for the most appropriate threshold.

We conclude this section by mentioning that the new method may be applicable to other microwave channels as well. Specifically, it is probably quite applicable to MHS Channel 4 at 183 ± 3 GHz, using SEVIRI Channel 6 at 7.3 μm for the matchup, since that channel combination also has roughly similar Jacobians, like the MHS 3/SEVIRI 5 combination used here. However, it is much more doubtful whether a similar method could be used for channels with surface influence, since such channels behave quite differently in the infrared and the microwave, due to the very different emissivity. On the other hand, for channels with surface influence, the dynamical range of ordinary SNOs is much higher, so there is also less need for the new matchup method.

8. Summary and Conclusions

Opportunistic constant target matching is a new method for satellite intercalibration. It is complimentary to the traditional SNO method because it can provide warm matchups in cases where the SNO method provides only cold matchups.

A geostationary infrared sensor is used to select constant target matches for two different microwave sensors. The main conditions for the method to work are (a) that the geostationary sensor is stable over the matchup time, which can be between several hours and many days, and (b) that the natural variability is much stronger than the diurnal cycle, so that the atmospheric variability inside the brightness temperature matchup threshold for the geostationary sensor is unbiased. We have good reasons to believe that both conditions are reasonably met in the selected case study, although this is difficult to quantify.

We explore the statistical properties of the new method with a simple Monte Carlo simulation. Real observations are found to behave consistently with these simulations, increasing our confidence that the method can be a valuable tool for intercalibration efforts.

Due to the temporal stability requirement, the new method effectively screens out high clouds, since they typically are more variable than the clear atmosphere. Furthermore, we apply the method only in the tropics, so there are no cold surface cases. These factors combine, with the consequence that the new method does not generate any very cold matchups, as shown in Figure 1. This means that the traditional SNO method is still needed additionally in order to cover the full dynamic range.

The large dynamic range of the new OCTM and the traditional SNO method combined can allow for future studies of the radiance dependence of the bias. In the examples in this article, we have for simplicity assumed the bias to be a simple offset, but this will likely not be the case for some sensors. This is one reason why it is so desirable to cover the full dynamical range of the radiances with matchups.

Acknowledgments
The Python script with the Monte Carlo simulation, as well as the matchup data used to generate Figures 5 and 1, is available in the Zenodo data repository (doi: https://doi.org/10.5281/zenodo.3368988). The SEVIRI data are Level 1.5 data provided by EUMETSAT. The MHS data are Level 1c data processed with AAPP, from Level 1b obtained from EUMETSAT. The time range is months October to December 2013. Both daytime and nighttime data were used in the study indiscriminately. We gratefully acknowledge the support of the Typhon community, in particular Gerrit Holl and Lukas Kluft, and the ARTS community, in particular Oleksandr Bobryshev for the MHS setup and Max Schaper for the SEVIRI setup. Our interest in this issue was triggered by the FIDUCEO project (https://www.fiduceo.eu), lead by Chris Merchant, which also partially supported M. P. and M. B. (Grant Agreement 638822). This study also contributes to Excellence Cluster "Climate, Climatic Change, and Society"—Project 390683824, and to the Center for Earth System Research and Sustainability (CEN) of Universität Hamburg. During the finalization of this article, M. P. was funded by the Deutsche Forschungsgemeinschaft (DFG) in Project "Elevated Moist Layers," Grant BU 2253/9-1, as part of the HALO research program.

References

Anderson, G. P., Keizys, F. X., Chetwynd, J. H., Wang, J., Hoke, M. L., Rothman, L. S., et al. (1995). FASCODE/MODTRAN/LOWTRAN: Past/present/future. In *18th Annual Review Conference on Atmospheric Transmission Models*.

Bentley, J. L., & Maurer, H. A. (1980). Efficient worst-case data structures for range searching. *Acta Informatica*, *13*(2), 155–168.

Bonsignori, R. (2007). The Microwave Humidity Sounder (MHS): in-orbit performance assessment. In S. Habib, R. Meynart, S. P. Neeck, & H. Shimoda (Eds.), *Sensors, systems, and next-generation satellites xi* (Vol. 6744, pp. 93–104). SPIE. https://doi.org/10.1117/12.737986

Buehler, S. A., Mendrok, J., Eriksson, P., Perrin, A., Larsson, R., & Lemke, O. (2018). ARTS, the atmospheric radiative transfer simulator—Version 2.2, the planetary toolbox edition. *Geoscientific Model Development*, *11*(4), 1537–1556. https://doi.org/10.5194/gmd-11-1537-2018

Buehler, S. A., Östman, S., Melsheimer, C., Holl, G., Eliasson, S., John, V. O., et al. (2012). A multi-instrument comparison of integrated water vapour measurements at a high latitude site. *Atmospheric Chemistry and Physics*, *12*(22), 10,925–10,943. https://doi.org/10.5194/acp-12-10925-2012

Cao, C., Weinreb, M., & Xu, H. (2004). Predicting simultaneous nadir overpasses among polar-orbiting meteorological satellites for the intersatellite calibration of radiometers. *Journal of Atmospheric and Oceanic Technology*, *21*, 537–542. https://doi.org/10.1175/1520-0426(2004)021<0537:PSNOAP>2.0.CO;2

Cao, Ch., Xu, H., Sullivan, J., McMillin, L., Ciren, P., & Hou, Y.-T. (2005). Intersatellite radiance biases for the High-Resolution Infrared Radiation Sounders (HIRS) on board NOAA-15,-16, and-17 from simultaneous nadir observations. *Journal of Atmospheric and Oceanic Technology*, *22*(4), 381–395.

Chung, E.-S., Soden, B. J., & John, V. O. (2013). Intercalibrating microwave satellite observations for monitoring long-term variations in upper and mid-tropospheric water vapor. *Journal of Atmospheric and Oceanic Technology*, *30*, 2303–2319. https://doi.org/10.1175/JTECH-D-13-00001.1

Chung, E. S., Soden, B. J., Sohn, B. J., & Schmetz, J. (2013). An assessment of the diurnal variation of upper tropospheric humidity in reanalysis datasets. *Journal of Geophysical Research: Atmospheres*, *118*, 3425–3430. https://doi.org/10.1002/jgrd.50345

Chung, E. S., Sohn, B. J., Schmetz, J., & Koenig, M. (2007). Diurnal variation of upper tropospheric humidity and its relations to convective activities over tropical Africa. *Atmospheric Chemistry and Physics*, *7*(10), 2489–2502. https://doi.org/10.5194/acp-7-2489-2007

Dolatshah, M., Hadian, A., & Minaei-Bidgoli, B. (2015). Ball*-tree: Efficient spatial indexing for constrained nearest-neighbor search in metric spaces. arXiv.org.

EUMETSAT (2019). Typical radiometric noise, calibration bias and stability for Meteosat-8, -9, -10 and -11 SEVIRI (Tech. Rep.). Eumetsat-Allee 1, D-64295 Darmstadt, Germany: EUMETSAT. (*Doc.No.:EUM/OPS/TEN/07/0314*).

Hans, I., Burgdorf, M., Buehler, S. A., Prange, M., Lang, T., & John, V. O. (2019). An uncertainty quantified fundamental climate data record for microwave humidity sounders. *Remote Sensing*, *11*(5), 548. https://doi.org/10.3390/rs11050548

Hewison, T. J., Wu, X., Yu, F., Tahara, Y., Hu, X., Kim, D., & Koenig, M. (2013). GSICS inter-calibration of infrared channels of geostationary imagers using Metop/IASI. *IEEE Transactions on Geoscience and Remote*, *51*(3), 1160–1170.

Holl, G., Buehler, S. A., Rydberg, B., & Jiménez, C. (2010). Collocating satellite-based radar and radiometer measurements—Methodology and usage examples. *Atmospheric Measurement Techniques*, *3*, 693–708. https://doi.org/10.5194/amt-3-693-2010

Iacovazzi, R. A. Jr., & Cao, C. (2007). Quantifying EOS aqua and NOAA POES AMSU-A brightness temperature biases for weather and climate applications utilizing the SNO method. *Journal of Atmospheric and Oceanic Technology*, *24*, 1895–1909. https://doi.org/10.1175/JTECH2095.1

John, V. O., Allan, R. P., Bell, B., Buehler, S. A., & Kottayil, A. (2013). Assessment of inter-calibration methods for satellite microwave humidity sounders. *Journal of Geophysical Research: Atmospheres*, *118*, 4906–4918. https://doi.org/10.1002/jgrd.50358

John, V. O., Holl, G., Atkinson, N., & Buehler, S. A. (2013). Monitoring scan asymmetry of microwave humidity sounding channels using Simultaneous All Angle Collocations (SAACs). *Journal of Geophysical Research: Atmospheres*, *118*, 1536–1545. https://doi.org/10.1002/jgrd.50154

John, V. O., Holl, G., Buehler, S. A., Candy, B., Saunders, R. W., & Parker, D. E. (2012). Understanding inter-satellite biases of microwave humidity sounders using global simultaneous nadir overpasses. *Journal of Geophysical Research*, *117*, D02305. https://doi.org/10.1029/2011JD016349

Kottayil, A., John, V. O., & Buehler, S. A. (2013). Correcting diurnal cycle aliasing in satellite microwave humidity sounder measurements. *Journal of Geophysical Research: Atmospheres*, *118*, 101–113. https://doi.org/10.1029/2012JD018545

Kottayil, A., John, V. O., Buehler, S. A., & Mohanakumar, K. (2016). Evaluating the diurnal cycle of upper tropospheric humidity in two different climate models using satellite observations. *Remote Sensing*, *8*(4), 325. https://doi.org/10.3390/rs8040325

Mears, C. A., & Wentz, F. J. (2005). The effect of diurnal correction on satellite-derived lower tropospheric temperature. *Science*, *309*(5740), 1548–1551. https://doi.org/10.1126/science.1114772

Mrziglod, J. (2018). Using collocated satellite data for ice water path retrieval—A reimplementation of SPARE-ICE (Master's Thesis), Universität Hamburg, Fachbereich Geowissenschaften, Meteorologisches Institut.

Nagle, F. W., & Holz, R. E. (2009). Computationally efficient methods of collocating satellite, aircraft, and ground observations. *Journal of Atmospheric and Oceanic Technology*, *26*, 1585–1595.

Ohring, G., Wielicki, B., Spencer, R., Emery, B., & Datla, R. (2005). Satellite instrument calibration for measuring global climate change. *Bulletin of the American Meteorological Society*, *86*, 1303–1313. https://doi.org/10.1175/BAMS-86-9-1303

Saunders, R. W., Blackmore, T. A., Candy, B., Francis, P. N., & Hewison, T. J. (2013). Monitoring satellite radiance biases using NWP models. *IEEE Transactions on Geoscience and Remote*, *51*(3), 1124–1138. https://doi.org/10.1109/TGRS.2012.2229283

Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., & Ratier, A. (2002). An introduction to Meteosat Second Generation (MSG). *Bulletin of the American Meteorological Society*, *83*(7), 977–992.

Scott, N. A., Chedin, A., Armante, R., Francis, J., Stubenrauch, C. J., Chaboureau, J.-P., et al. (1999). Characteristics of the TOVS pathfinder Path-B dataset. *Bulletin of the American Meteorological Society*, *12*, 2679–2701.

Shi, L., & Bates, J. J. (2011). Three decades of intersatellite-calibrated High-Resolution Infrared Radiation Sounder upper tropospheric water vapor. *Journal of Geophysical Research*, *116*, D04108. https://doi.org/10.1029/2010JD014847

Shi, L., Bates, J. J., & Cao, C. (2008). Scene radiance-dependent intersatellite biases of HIRS longwave channels. *Journal of Atmospheric and Oceanic Technology*, *25*(12), 2219–2229.

Tabata, T., John, V. O., Roebeling, R. A., Hewison, T., & Schulz, J. (2019). Recalibration of over 35 years of infrared and water vapor channel radiances of the JMA geostationary satellites. *Radio Science*, *11*, 1189. https://doi.org/10.3390/rs11101189

Wang, L., Cao, C., & Ciren, P. (2007). Assessing noaa-16 hirs radiance accuracy using simultaneous nadir overpass observations from AIRS. *Journal of Atmospheric and Oceanic Technology*, *24*, 1546–1561. https://doi.org/10.1175/JTECH2073.1

Wang, L., Wu, X., Goldberg, M., Cao, C., Li, Y., & Sohn, S.-H. (2010). Comparison of AIRS and IASI radiances using GOES imagers as transfer radiometers toward climate data records. *Journal of Applied Meteorology and Climatology*, *49*(3), 478–492. https://doi.org/10.1175/2009JAMC2218.1

Yu, F., Wu, X., Rama Varma Raja, M. K., Li, Y., Wang, L., & Goldberg, M. (2013). Diurnal and scan angle variations in the calibration of GOES Imager infrared channels. *IEEE Transactions on Geoscience and Remote Sensing*, *51*(1), 671–683. https://doi.org/10.1109/TGRS.2012.2197627

Zou, C.-Z., Goldberg, M. D., Cheng, Z., Grody, N. C., Sullivan, J. T., & Tarpley, D. (2006). Recalibration of microwave sounding unit for climate studies using simultaneous nadir overpasse. *Journal of Geophysical Research*, *111*, D19114. https://doi.org/10.1029/2005JD006798

Zou, C.-Z., & Wang, W. (2011). Intersatellite calibration of AMSU-A observations for weather and climate applications. *Journal of Geophysical Research*, *116*, D23113. https://doi.org/10.1029/2011JD016205