Dear authors, thanks a lot for considering my comments. The manuscript has been significantly improved. However, I have a major concern regarding formulas the authors give. In addition, I have some minor comments. Please find my comments below. The author's latest comments are in blue color. My new comments are in green.

Author's response: We agree about the need of characterizing the target to provide final uncertainty results. However, we do not agree that the lack of the target characterization cancels the validity of the results, since we used a theoretical model of target RCS that has all the properties that a calibrated target would have, except the absolute values.
The results presented in the article enable the identification of several uncertainty sources, as well as their relative contribution to the experiment uncertainty. This information can be very valuable to design future calibration experiments based on reference reflectors, whether they are mounted on masts or held by other means, such as UAVs. The underlying principles remain the same.
We also add that our objective is not to claim we have a reference instrument, but to present all the information and advancements obtained from our experimental campaigns, specially in uncertainty characterization. For example, with the results we can quantitatively compare two different experimental setups, finding different factors limiting uncertainty for each (SCR for the 10 m mast, alignment for the 20 m mast). Additionally, we are not aware of any other published methodology of radar calibration that considers the bias introduced due to misalignment between target and radar. For us this work is a step towards more precise calibration methodologies, and we expect it to act as a reference to improve the preparation of future calibration experiments. Because it is true that at this stage we can only do a rough estimation of RCS uncertainty, we agree to highlight this explicitly.
This is now stated/included in:
Line 14 of the Abstract.
Lines 197-200
Line 229
Lines 507-509
Calibration result for all experiments (lines 512-531)
Table 2
Lines 585-587
This also implied modifications to some text in the article to remain consistent:
Lines 18-20 of the abstract.
Lines 54-55 of the Introduction
Lines 153-154
We also added an estimation of the maximum uncertainty in RCS characterization required to reach a calibration uncertainty of 0.5 dB in lines 592-597.

Comment: Thanks for these changes. The modifications you have introduced make it much clearer for a reader that the main source of uncertainties is the characterization of the used reflectors and that this is not covered in this study. Just a few minor comments here:

- Please be consistent with the goal value. It is 1 dB in the abstract (line 20) but 0.5 dB throughout the text (e.g. lines 55, 593).
- In lines 593 – 594 you apply the formula of the std of a sum of two uncorrelated variables, if I understand it right (sqrt(0.4^2+0.3^2) = 0.5). I would agree with it if one would use a newly manufactured reflector every time, but most likely, it will be just a single one so the bias due to the corner reflector will be constant (zero variance, the formula is not applicable). In this case,

the reflector contributes to the systematic error, while the effects considered in the manuscript characterize the random error. I would recommend making this clear.

- Since the study makes conclusions on the total uncertainties with a number of assumptions (i.e. parallel antennas, flat noise figure, reflector with known cross section), please consider a change of the title of the manuscript to something like "Aspects of cloud radar calibration based on corner reflectors".

Author's response: Following this suggestion, we now present the received power units to dBm and the calibration terms to dB. This led to the recalculation of the calibration terms absolute value.
Modifications:
- Because of this improvement, lines 79-80 and 82-84 explaining these arbitrary power units are no longer necessary and were removed.
- Eq. (2a) and line 542 were modified to remove the unnecessary digital gain term kd.- Writing of the units corrected in lines 97, 112, 127, 142 and the Glossary.
- Absolute value of power/calibration terms corrected in lines 261, 264, 265, 267, 319-321, Table 1, and lines 512-531 of the calibration results.
- Additionally, Figures 3, 4(a), 5, 6 and 10 were modified to remain consistent with the power units.

Comment: Here I still have a major concern. I would like to thank the authors for their efforts, but the equation 2 is still wrong. It cannot give dB units, because the ratio in the parenthesis is still not unitless. Please note, that this is already the third time I ask to adjust all the terms properly.

I try to do my best to explain what I expect if I were using the method proposed by the authors. Typing long formulas in Word is a bit inconvenient. Therefore, I wrote my considerations on paper.

$P_r = P_m \cdot C_r$ (Eq. 1)

P_r - real (expected from the radar equation) received power [W]

P_m - measured received power [W]

C_r - calibration term [unitless]

In log scale:

$P_r\,[dBm] = P_m\,[dBm] + C_r\,[dB]$ (Eq. 2)

In the case of an ideal radar we expect:

$$P_r = \frac{\pi^3}{16 \log_e 2} \cdot \frac{P_t \cdot h}{\lambda^2} \cdot G_t \cdot G_r \cdot \theta^2 \frac{1}{R^2} |K|^2 Z \quad \text{(Eq. 3)}$$

Eq. 3 is for meteorological targets and is taken from Probert - Jones 1964.

P_t - transmitted power [W]

h - range resolution [m]

λ - wavelength [m]

G - antenna gain [unitless]

R - range [m]

$|K|^2$ - dielectric factor [unitless]

Z - reflectivity [$mm^6 m^{-3}$]

$$P_m = P_r \cdot \underbrace{\frac{L_t}{L_r} \cdot \frac{G_t^{real}}{G_t^{assum}} \cdot \frac{G_r^{real}}{G_r^{assum}} \cdot \frac{\theta^2_{real}}{\theta^2_{assum}} \cdot \frac{1}{L_o}}_{\text{estimated in the study}} \cdot \overbrace{\frac{1}{L_{\parallel}} \cdot \frac{L_{at}^{real}}{L_{at}^{assum}} \cdot \frac{Z^{real}}{Z^{assum}}}^{\text{not estimated}} \quad \text{(Eq. 4)}$$

C_r (unitless)

L_t - losses in transmitter path

L_r - losses in receiver path

L_o - losses due to the antenna overlap

$L_{\parallel}$ - losses in the case antennas are not parallel

In the case of a point target $\frac{Z^{real}}{Z^{assum}}$ should be $\frac{\Gamma^{real}}{\Gamma^{assum}}$.

where Γ is the reflector cross-section in m^2

Indecks 'real' and 'assum' correspond to real (true) values and assumed ones, respectively.

I forgot to write one term. Lat is the attenuation due to propagation to the target and back.

In my point of view, the authors should use the calibration term Cgamma as in Eq. 4 in the drawing. It is clear how to use it in order to correct the measurements (it works in the same way for power and reflectivity values) and it is unitless as it should be. I marked two components of the calibration term, the one characterized in the manuscript (greed rectangle) and the one with certain assumptions (red rectangle). I believe such a separation in the very beginning of the manuscript would help a reader to understand which effects are considered and which are not.

Author's response: We originally used range because BASTA-Mini has only 4 standard operational modes, so it was not a very important distinction. Nevertheless, we agree that this formulation would improve the generality of the method, and consequently we modified all range dependent terms by terms depending on the IF frequency F_b.
Besides, we added a short explanation of the F_b term in lines 75-81 and the new Equation (1), which indicates how F_b it is associated with the range r.

Comment: Thanks, it is better now. The method in general can be used for other FMCW radars as well and a reader should understand that the receiver loss depends on IF and not on range.

Author's response: In this section we do not intend to estimate the absolute value of gain at the IF, but rather to quantify relative gain changes with respect to the calibrated IF frequency (associated with the target position). To do this using passive observations, we had to make the assumption of a constant noise power, both from the system and from environment in the 12 MHz bandwidth of the receiver. This assumption is reasonable because components used in the receiver have a much larger bandwidth (for example for the Low Noise Amplifier (LNA) it is of 35 GHz).
We indicated briefly this assumption in the previous version, but we didn't mention the impact it could have on uncertainty. This is now estimated from the LNA specifications. Ist variability of gain and noise figure in the 12 MHz bandwidth used is smaller than 0.1 dB. Since LNA are typically the main source of system noise in the receiver, we consider that 0.1 dB is a safe estimation of the uncertainty introduced by assuming a constant system noise in the IF bandwidth. This term dominates the RMSE between the fit and data, and the inter-period variability, thus we now define the IF correction function uncertainty to be of 0.1 dB.
Despite this, we agree that two point calibrations are highly desirable because they enable the retrieval of receiver absolute gain and system noise. This is now indicated in the text. Thus, article changes are:
- More accurate explanation of passive observations (lines 365-368, equation 9a)
- Brief indication of the benefits of performing two point calibrations for receivers (lines 370-373)
- Explanation of the constant system noise assumption and introduced uncertainty (lines 374-383, equation 9b)
- Clarification in the explanation on how the IF correction function is retrieved (lines 384-396, line 406-408).
- Final uncertainty of the IF correction function (lines 413-414)

Comment: A couple of major concerns:

- Lines 265-267: formulation is wrong. The very first component in the receiver chain is a low noise **amplifier**. The receiver amplifies the received signal not reduces the power level. Despite on the amplification LNA reduces the signal-to-noise ratio and this is characterized by its noise figure. I recommend to use the formula I gave last time for the proper explanation.

- Formula 9: If noise powers are in linear units, why dB values of losses are subtracted? If noise powers are in dBm they cannot be summed.

Also a minor concern:

- Line 376: please be careful here. I completely agree that noise figure of LNA can be assumed flat. Since noise figure of mixers, active IF filters and ADC units can be very high, these components can still contribute to the noise figure despite the amplification of LNA. The question is how about other components and standing waves, which can produce wavy shape (> 100 K variability) of the system noise temperature even within several MHz bandwidth? If this was considered please mention, if not please explain that other effects caused by other components of the receiving chain or standing waves can affect the assumption of the constant noise temperature.

- Taking into account the three comments, I recommend a revisiting the section 5.5 lines 364 - 384.

Author's response: Thank you for this proposal, we believe it is a very good idea with good potential. Because of this, we performed several theoretical calculations to check if we could estimate a range of possible antenna misalignment angles and the associated uncertainty with our data. Summarizing, our results show that the experimental setup used is not appropriate for this measurement, but they also indicate us a path to perform such experiments in the future. Since the targets used at 196 and 376.5 are different, the uncertainty in the **calibration coefficient difference** at both distances is very large (~3 dB). This uncertainty makes it impossible to bound the possible alignment within 1.5 degrees, which is our antenna characterization width. This large uncertainty comes mostly from the use of two different calibration targets. This decision was made because the experiment was designed to applicability of the absolute calibration method for different experimental setups, and because the proposed experiment was not considered at the time. Given that the proposed experiment was not done during the campaigns presented in the article, we have no way to gather any additional information on antenna alignment for that period. Thus, we leave this section unchanged with respect to the previous version. Yet, with the theoretical calculations we found that if we get an uncertainty in the order of 0.5 dB when comparing calibration constants at these two different distances, antenna misalignment could be constrained to values ranging in the order of tenths of degree. This could be achieved, for example, by using the 20 meter mast setup at both distances using the same reflector each time. It is worth noting that the tools developed for this analysis now enable the design of an experiment with optimized parameters for this retrieval. Taking all this into consideration, in our opinion the potential of the proposed experiment indicate that it must be further studied for its implementation in future calibration campaigns.

Comment: I understand the point of the authors. I would recommend to do the following. Please summarize the effects which you have characterized in your study and write that the uncertainty you have found are only related to these effects. Also specify explicitly a list of effects to be characterized in the future (target, IF dependency of noise figure, parallelism of antennas, etc).

Author's response: This happens because, to capture the widest range of possible temperatures, we had to use longer time series of data. The use of longer time series introduced some points measured under suboptimal conditions. For example, we have observed that high wind speeds lead to larger variabilities in the calibration value due to oscillations of both the mast and the radar. Meanwhile, drizzle adds a time dependent bias, most likely caused by changes in wet radome (and wet target) attenuation over time. An effort was done to clean the dataset, but inevitably some noisy data points remained. This is now stated more clearly in lines 345 to 348. Therefore, to estimate the temperature correction function and its uncertainty we did an statistical analysis of data, and then used the RMSE between the model and data points as the estimation of model uncertainty. To bound the uncertainty value we calculated the RMSE of the model for each degree of deviation from the reference temperature, obtaining a RMSE range between 0.07 to 0.23 for 0 and +3 degrees of deviation respectively. We also checked the bias per degree for each iteration and for all the dataset, and found out that its mean value is always within +- 0.2 dB with respect to the model. Therefore, we now state that the temperature correction function uncertainty is less or equal to 0.23. This change is reflected in lines 353-357. This is also true for the mentioned case, were most points are covered by the other iterations data. The larger spread reaching deviations of 0.5 dB in this case are caused by short period of drizzle that happened in this iteration. However, since this data, and the rest of data that deviates from the model is also included in the calculation of RMSE, we think this is a reliable criteria for the estimation of the temperature correction function uncertainty.

Comment: Thanks for the explanation. It is clear now.