Dear Dr. Alexander Myagkov,

Thank you for this careful and detailed review. Your comments were very valuable to improve the quality of the paper and the results. We provide here a point-by-point response to address each review comment (shown in blue color). Our responses appear in black color. In addition we explain all changes made to the manuscript. They can be seen in the manuscript document that includes change-tracking, attached to this file.

Additionally, we found that the reviewers comments were very helpful to improve the quality of our manuscript. Both reviewers are acknowledged in the acknowledgment section.

1. Line 113: I would expect that the calibration targets are significant and actually the main source of uncertainties, which is not discussed in the manuscript. RCS are given up to the 4th significant digit. What are the uncertainties in these values? Were there any measurements of the targets made in an anechoic chamber? From a book about the radar reflectors, I know that the manufacturing precision is super critical for the RCS of corner reflectors. In the case of reflectors, which are much larger than the wavelengths, one degree deviation from perpendicularity of the reflector planes causes a change in RCS of about 10 dB. Since the aim of the proposed calibration procedure is to reach the calibration in the order of a few 0.1 dB (such an error could be easily caused by a tiny inperpendicularity), it would be very helpful to know how the authors evaluated this and what the uncertainty is. A reference with relevant information:

- **Section 2.4 in Garthwaite et al 2015. The Design of Radar Corner Reflectors for the Australian Geophysical Observing System. (and references therein)**

The main objective of the paper is to present a calibration methodology. The methodology itself is not affected by the use of a real or theoretical target RCS. Actually, once the real target RCS is retrieved, any possible bias in the results can be corrected without changing the calibration method. We now state this more clearly in lines 190-196. The company that manufactures the targets declares having a cutting accuracy better than 0.1 mm and an alignment precision better than 0.1°, therefore we can expect a bias but it should be on the order of 1-2 dBsm.

We also include now how to account for the uncertainty of an eventual target characterization (eq. 6a and lines 231-234), and indicate that the uncertainty of the target calibration may increase the uncertainty in the results (lines 530-535).

Finally, as future work we now include the need of a target characterization in an anechoic chamber to correct any bias introduced by the use of the theoretical model (lines 598-600).

2. I strongly recommend to carefully check the units throughout the manuscript. Often the units do not fit. These are just occasions I noticed, there could be other.

a. **In the line 61 it is mentioned that Cgamma(T) is in dB. When I look into the Eq. 1a where this parameter is calculated I see a dimensionless number in the numerator and m^2*W in the denominator. 10log gives dB only when the ratio is dimensionless.**
b. **Line 77: what does unreferenced dB mean? Power shall be in W, mW, dBm or something similar. It is not a unitless ratio and thus cannot be in just dB.**
c. **Lines: 195 – 200: Power is in dB. Please see the comment 1b.**
d. **Lines: 237 – 241: Power is in dB. Please see the comment 1b.**
e. **Figures 3, 4, power is in dB. Please see the comment 1b.**
f. **RCS has different units throughout the manuscript. In line 113 it is in m^2, line 269 dBsm (is it dB related to square meter?), in Fig. 6 it is in dB (again cannot be because it is not an unitless ratio).**

The problem with the power units arises because power output in the BASTA radar is in an arbitrary power unit. We define this power unit as **dB(AU) = 10log$_{10}$(AU)**. The arbitrary unit defined as **AU** is proportional to watts multiplied by a unitless digital gain **k_d**, which depends on the digital signal processing configuration of the radar, such that **dB(AU) = dBW + 10log$_{10}$(k_d)**. Since the absolute calibration method will provide a calibration result that compensates this constant term, we did not work in transforming the power to standard physical units.

We now explained this detail in lines 72-76 . For consistency, now every power unit is defined in dB(AU) units, and therefore the RCS calibration is now in dB(AU^{-1} m^{-2}) and the reflectivity calibration is in dB(mm^{-6} m^{-5} AU^{-1}). This way, when the term is multiplied by reflected power and distance to the corresponding power, the result will be in the correct units (dBsm or dBZ).

All RCS values presented in the manuscript are now in dBsm units, both in text and figures. Line 84 also indicates that dBsm units are decibels referenced to a square meter. We also fixed a typo in Fig. 9 (prev fig 6). The maximum RCS indicated before in the label was of 28.28 dBsm, but it is actually 28.34 dBsm.

3. Line 144: I would also recommend to mention which FFT window have been used for ranging, since it defines how many range bins to sum.

During calibration we used a Hann time window, which is the default for the BASTA radar. This is now mentioned in line 178. Additionally, we include a new figure (Figure 3) to show which gates are used to estimate the target signal. The integration of additional gates increases the signal power by less than 0.01 dB, as indicated in lines 183-185.

4. Since FMCW radars have intermediate frequency (IF) filters, the calibration can differ for different range bins due to different filter gains at different IF frequencies. How the frequency response of the IF filter is taken into account in the calibration method?

Many thanks for this observation. Indeed we did not consider gain variations in the radar IF in the methodology. This comment motivated a significant improvement in the proposed Calibration Methodology, which now considers range variability in the receiver loss in addition to temperature effects. Because changes were important, they are introduced in several sections:

Abstract: included the IF correction in lines 8-9

Section 2: Equations used in Radar Calibration
This change is included in every mention of receiver losses as $L_r(T, r)$, and is therefore propagated to the RCS and reflectivity calibration terms as well, which now depend on temperature and range ($C_\Gamma(T, r)$, $C_Z(T, r)$).
Lines 103-111 and equation (2) present how the IF correction is taken into account. Lines 114 to 119 explain in which order the calibration calculation is performed (first calibration at the mast position, then application of the IF correction).

Section 4: Methodology
Figure 2 now includes the IF correction as a block necessary for radar calibration. Lines 204-206 introduce how this correction function is retrieved.
Line 217 and eq. 6a now includes an IF correction uncertainty term.

Section 5.5: IF loss correction function
This is a new section explaining how we retrieve the IF correction function.

Section 6: Results
The results section now includes the value of the IF correction uncertainty and the IF correction functions retrieved for each calibration experiment in Table 2.

Section 7: Conclusions
The study of IF losses as a necessary part of calibration is now mentioned in lines 577-580.

5. Correct me if I am wrong but as far as I understand C^0_gamma in lines 153-155 is different from the one calculated in Eq 4. In these lines C^0_gamma is calculated from each sample within an iteration, while the one in Eq. 4 is calculated from averages of several iterations.

Yes, you understood correctly. To improve readability we modified the name of the C_Γ^0 of a single sample within an iteration to $C_{\Gamma s}^0$ and added a short explanation, in lines 187-188.

6. Line 209: As far as I understand the lab pattern (Fig. 3b) characterizes only one antenna. In Fig. 3b I see a very high variability of the measurements (green and yellow dots). I do not see how based on these result the conclusion that the two antennas are parallel can be made.

The dots are the power received from the target normalized with respect to the maximum value and divided by 2, because we assume antennas are identical. This allows a comparison between data from scans and the laboratory antenna pattern. This is now explained in the caption of figure 4b (old 3b).

However, we did another revision of the scanning data and concluded that, at present, it is not possible to retrieve alignment information with an accuracy comparable to the antenna beam-width. This is now stated in lines 294-295. The reason is that the repeatability of the scanner positioning is not sufficient to allow a reliable retrieval under our current procedure.

Additionally, we now include a discussion on how parallax errors can influence the measurements (286-290), and indicate that calibration results are compatible with parallax errors smaller than the radar beamwidth (296-298). Since we don't have information on the exact alignment, we now mention the parallel antennas only as an hypothesis (245, 299-300).

Finally, we improved the calibration methodology by indicating how parallax errors can be taken into account, suggesting the addition of an additional range dependent correction function (300-301), and by introducing an uncertainty term representing the error in the antennas alignment estimation (eq. 6b, lines 243-245 and 301-302).

7. Line 265: when I look into fig 5d I see that for some entire iterations the measurement points are not within 0.13 dB from the black line. Some iterations (like red one and violet one) have deviation exceeding 0.13 dB as well. Is the given uncertainty sigma_T = 0.13 dB reliable in this sense?

To verify if data did follow a linear relationship, we did a new plot with the point density of all samples together. This figure has been added to the paper (Figure 7). In this figure it is easier to observe that deviated points are rather exceptional, with most points close to the regression. From this figure we think the 0.13 dB RMSE value is representative for most samples.

We also modified Figure 6 (D). Now it is only used to introduce the data set, with the linear fit shown in new Figure 7. This produced text changes in lines 351-355, and 360-371.

8. Are the numbers, given in line 311, resolution of the stepper or real angular accuracy? This can be tested if the same target angular position is reach from opposite rotation directions. Would the received power be the same?

This is the stepper resolution. Since the pointing algorithm relies on maximizing received power by scanning around the target, it is reasonable to assume that the uncertainty should be half of the stepper resolution. We agree that real angular accuracy is likely to be lower, as can be inferred from antenna scanning where received power variability is significant.

In any case, the RCS bias estimation of Figure 9 is just intended as an example to explain what is the misalignment bias and to show that its impact will most likely have a privileged direction.

The actual estimation of the bias correction is shown in section S3 of the supplementary material. This estimation relies on a sampling of all RCS distributions for randomly generated uncertainty sets. Specifically, for radar aiming we consider uncertainties from 0º to 0.375º, which is three times the stepper resolution. We improved the explanation of this procedure in lines 460-472 and in the supplementary material.

9. Fig. 7a: the sharp edges on the left side of the blue curve and on the right side of the yellow curve look so much different from the rest of the curves in this figure. What causes such an effect?

This observation is very important. There was in fact a problem with the data used for the plot. When calculating the calibration results we selected the data from the contiguous hour with lowest variability, but the plot was incorrectly made using all the data from each target sampling iteration, as in the temperature characterization experiment.
We also took this correction as an opportunity to also improve the explanation of the methodology, adding lines 169-170 to indicate more precisely how data is chosen for the calibration estimation, and lines 200-203 to indicate that the temperature correction estimation is done using longer sampling times of the target signal.

The sharp lines corresponded to the beginning and end of the sampling period in cases where there was a systematic drift in the calibration constant value (for example a period under a fast temperature change in one direction).

The correction of this bug didn't affect the calibration values estimated for the 20 m mast, as they were calculated using the correct dataset, but did introduce a change of 0.01 dB in the 10 m mast retrievals. This is corrected in table 1 to guarantee consistency.

10. Fig. 7a: The standard deviation of the yellow curve and the red one are close (0.09 and 0.13 respectively). But in the figure I see that the red curve is at least by a factor of 2 broader. Please double check the numbers.

This was caused by the same problem explained in comment 9.

11. From the table 2 I see that the uncertainty in the median values of the experiments A and B are < 0.03. 7a shows that the mean/median values for each iteration are distributed from -275.5 to -274.7. While 7b shows that all the iterations are in range of -274.2+/-0.2. If the number 0.03 is calculated for each iteration separately, which uncertainty component in the table 2 contains +/- 0.5 dB variability of the mean in the experiment A?

Since the effective RCS distribution for our system alignment is not gaussian, we couldn't use the standard deviation between iterations as a direct estimator of uncertainty.

Rather, what we did was to perform an estimation of the effective RCS distribution for our system using the observed variability between calibration results as input information. This effective RCS distribution is used to compute the most likely calibration value consistent with our observed spread through the estimator of Eq. (5), and therefore the uncertainty of this procedure will be represented by the uncertainty in each term of this estimator.

These terms are the uncertainty propagated through the calculation of the average calibration result ($\Sigma\sigma_i/N^2$ and σ_T/N) plus the uncertainty in the bias correction term (σ_Λ). This is now indicated explicitly in lines 468-472.

12. I think Lat in Eqs . 1b and 3b is also range dependent. Please indicate this.

We inserted this correction in the requested equations. We also corrected this same error in equation (7). Equations (1b) and (3b) were re-organized to correct for the ambiguity introduced when using both r and r0 (explained in the next point).

13. What is the difference between r0 and r in Eqs. 1b and 3b?

There was indeed an ambiguous use of r and r_0. This is corrected. Now r always represents the distance from the radar. It is used, for example, to indicate that a term is a function of distance.
Along this change, r_0 is now only used to indicate the distance of the target (from the radar). It is used for example to indicate that a certain variable dependent on distance is evaluated at the target position. This is now written in lines 93-94.
The distinction between r and r_0 is also explained in the newly added table of symbols.

14. Sometimes it is hard to follow the text because of a huge number of symbols. I therefore strongly recommend to add a table with a short description of all used symbols.

Thanks for this recommendation. We now include a table of symbols at the end of the manuscript to improve text readability.

15. In the Eq. 5 the authors assume that errors are not correlated from iteration to iteration. But if for example two consecutive iterations are made under similar conditions I would expect a certain correlation. In this case variances would not be reduced by factor of N^2 and N but by a smaller factor (Leith 1973, The standard error of time-average estimates of climatic means).

While developing the mathematical method for the estimation of the calibration coefficient C_r^0 , we made the underlying hypothesis that variations in the estimation of this value between different iterations are exclusively explained by an underlying probability distribution of effective RCS values. This distribution is generated by the alignment uncertainty of the experimental setup.

The method also includes the hypothesis that the resulting effective RCS after each system realignment is not correlated with the result of the previous one. As indicated in point 11 and as is now indicated in lines 460-472, by following this hypothesis we arrive to the calibration estimator of Eq. (5). During the derivation of this estimator, we get that the uncertainty is distributed between the terms $\Sigma\sigma_I/N^2$, σ_T/N and σ_Λ . The details of this derivation are in section S3 of the supplementary material.

To have meaningful results, we implemented many measures to have a system that behaves as close as possible to the theoretical model. Specifically, before each sampling iteration we purposely misalign the system very far from the operating conditions. This should reset the previous alignment state. Then, we follow the exact same alignment protocol each time, to ensure that the underlying distribution of the effective RCS remains unchanged for all iterations. This procedure is in fact a key aspect of the proposed calibration methodology.

Absolute Calibration method for FMCW Cloud Radars

Felipe Toledo[1], Julien Delanoë[2], Martial Haeffelin[3], Jean-Charles Dupont[4], and Susana Jorquera[2]

[1]Laboratoire de Météorologie Dynamique, École Polytechnique, Institut Polytechnique de Paris, 91128 Palaiseau, France
[2]LATMOS/IPSL, UVSQ Université Paris-Saclay, Sorbonne Université, CNRS, Guyancourt, France
[3]Institut Pierre Simon Laplace, École Polytechnique, CNRS, Institut Polytechnique de Paris, 91128 Palaiseau, France
[4]Institut Pierre-Simon Laplace, École Polytechnique, UVSQ, Université Paris-Saclay, 91128 Palaiseau, France

Correspondence: Felipe Toledo (ftoledo@lmd.polytechnique.fr)

Abstract.

This article presents a new Cloud Radar calibration methodology using solid reference reflectors mounted on masts, developed during two field experiments held in 2018 and 2019 at the SIRTA atmospheric observatory, located in Palaiseau, France, in the framework of the ACTRIS-2 research and innovation program.

The experimental setup includes 10 cm and 20 cm triangular trihedral targets installed at the top of 10 m and 20 m masts, respectively. The 10 cm target is mounted on a pan-tilt motor at the top of the 10 m mast to precisely align its boresight with the radar beam. Sources of calibration bias and uncertainty are identified and quantified. Specifically, this work assesses the impact of receiver compression, incomplete antenna overlap, temperature variations inside the radar, frequency dependent losses in the receiver IF, clutter and experimental setup misalignment. Setup misalignment is a source of bias previously undocumented in the literature, that can have an impact on the order of tenths of dB in calibration retrievals of W band Radars.

A detailed analysis enabled the design of a calibration methodology which can reach a cloud radar calibration uncertainty of 0.3 dB based on the equipment used in the experiment. Among different sources of uncertainty, the two largest terms are due to signal-to-clutter ratio and radar-to-target alignment. The analysis revealed that our 20 m mast setup with an approximate alignment approach is preferred to the 10 m mast setup with the motor-driven alignment system. The calibration uncertainty associated with signal-to-clutter ratio of the former is ten times smaller than for the latter.

Cloud radar calibration results are found to be repeatable when comparing results from a total of 18 independent tests. Once calibrated the cloud radar provides valid reflectivity values when sampling mid-tropospheric clouds. Thus we conclude that the method is repeatable and robust, and that the uncertainties are precisely characterized. The method can be implemented under different configurations as long as the proposed principles are respected. It could be extended to reference reflectors held by other lifting devices such as tethered balloons or unmanned aerial vehicles.

1 Introduction

Clouds remain to this day one of the major sources of uncertainty in future climate predictions (Boucher et al., 2013; Myhre et al., 2013; Mülmenstädt and Feingold, 2018). This arises partly from the wide range of scales involved in cloud systems, where a knowledge of cloud micro-physics, particularly cloud-aerosol interaction, is critical to predict large scale phenomena such as cloud radiative forcing or precipitation.

To address this and other related issues, the ACTRIS Aerosols, Cloud and Trace Gases Research Infrastructure is establishing an state of the art ground based observation network (Pappalardo, 2018). Within this organization, the Centre for Cloud Remote Sensing CCRES is in charge of creating and defining calibration and quality assurance protocols for the observation of Cloud properties across the complete network.

One of the key instruments for cloud remote sensing stations is the Cloud Radar. Cloud radars enable retrievals of several relevant parameters for cloud research, including but not limited to liquid water and ice content profiles, cloud boundaries, cloud fraction, precipitation rate and turbulence (Fox and Illingworth, 1997; Hogan et al., 2001; Wærsted et al., 2017; Dupont et al., 2018; Haynes et al., 2009). Additionally, recent studies revealed the potential of cloud radars to support a better understanding of fog processes (Dupont et al., 2012; Boers et al., 2013; Wærsted et al., 2019).

However, calibration remains a crucial factor in the reliability of radar retrieved data (Ewald et al., 2019). Systematic differences of 2 dB have already been observed, for example, between the satellite based radar CloudSat and the Lindenerg MIRA (Protat et al., 2009). This is a very important issue, since calibration errors as small as 1 dB dB would already introduce uncertainties in liquid water and ice content retrievals in the order of 15-20% (Fox and Illingworth, 1997; Ewald et al., 2019).

Since the objective of the CCRES is to guarantee a network of high quality observations, it is essential to develop standardized and repeatable calibration methods for its instrumental network.

This paper presents an absolute calibration method for W band radars. It has been developed based on results from two experimental calibration campaigns performed at the SIRTA Atmospheric Observatory, located in Palaiseau, France (Haeffelin et al., 2005). The SIRTA observatory hosts part of the ACTRIS CCRES infrastructure. For the experiments we used a BASTA-Mini W band Frequency Modulated Continuous Wave (FMCW) Radar, with scanning capabilities (Delanoë et al., 2016). Nevertheless, the principles, procedures and limitations presented here should be applicable for any radar with similar characteristics, even when operating in another frequency band.

The method consists on an end-to-end calibration approach, consisting in retrieving the radar calibration coefficient by sampling the power reflected from a reference reflector mounted on top of a mast (Chandrasekar et al., 2015). A depth detailed analysis of uncertainty and bias sources is performed, with the objective of reducing uncertainty under $0.5\ dB$. This low uncertainty in the calibration would not only be useful for high quality retrievals, but also enables the use of the radar as a reliable reference for calibration transfer to other ground or space based cloud radars (Bergada et al., 2001; Protat et al., 2011; Ewald et al., 2019).

The article is structured as follows: Section 2 present the ecuations and theoretical considerations involved in the calibration exercise. Section 3 shows the experimental setup, complemented by section 4 where the experimental procedure and data treatment is presented. Section 5 presents an analysis of the sources of uncertainty and bias involved in our calibration experiment. Section 6 presents the final calibration results, the uncertainty budget and an analysis of the variability in the calibration bias correction, followed by the conclusions.

2 Equations used in Radar Calibration

The absolute calibration of a radar consists in determining the calibration terms C_Γ and C_Z. They enable the calculation of Radar Cross Section Γ $\Gamma(r)$ (RCS) or Radar Equivalent Reflectivity Z_e Z_e respectively, from the power backscattered by a punctual or distributed target towards the radar (Bringi and Chandrasekar, 2001).

Equation (1a) presents an expression for the RCS calibration term $C_\Gamma(T)$ $C_\Gamma(T,r)$ of a FMCW radar , in dB, as a function of its internal parameters. The deduction of this expression is shown in the supplementary material. G_t and G_r are the maximum gain of the transmitting and receiving antennas respectively, dimensionless. λ λ is the wavelength of the carrier wave in meters and P_t p_t is the power emitted by the radar in watts.

The gain of solid state components changes with variations in their temperature T. Thus we make this dependence explicit in the receiver loss budget $L_r(T)$ $L_r(T,r)$ and in the transmitter loss budget $L_t(T)$. The loss budget is the product of all losses divided by the gain terms at the end of the receiver or emitter chain, and has no dimensions.

Additionally, a range dependence is included in $L_r(T,r)$ to account for variations in the receiver IF loss for different beat frequency values. The beat frequency is used in FMCW radars to calculate the distance between the radar and reflective targets (Delanoë et al., 2016). Changes in the IF loss for different beat frequencies introduce a range dependent loss.

The term k_d is included to account for the units of radar measured power. Radar measured power is in arbitrary power units defined as $dB(AU) = 10\log_{10}(AU)$.

These units are related to physical power units by the equation $dB(AU) = dBW + 10\log_{10}(k_d)$, where dBW is the physical power expressed in decibels relative to 1 watt and k_d is a unitless constant determined by the digital signal processing configuration.

In theory, $C_\Gamma(T)$ $C_\Gamma(T,r)$ can be calculated by characterizing the gains and losses of every component inside the radar system and adding them. This can be very challenging, depending on the complexity of the radar hardware and the available radio frequency analysis equipment. In addition, with this procedure it is not possible to quantify losses due to interactions between different components, specially changes in antenna alignment or radome degradation (Anagnostou et al., 2001). This motivates the implementation of an end-to-end calibration, which consists on the characterization of the complete radar system at once by using a reference reflector and Eq. (1b).

$$C_\Gamma(T,r) = 10\log_{10}\left(\frac{L_t(T)(4\pi)^3}{L_r(T)G_tG_r\lambda^2P_t}\;\frac{L_t(T)L_r(T,r)(4\pi)^3}{k_dG_tG_r\lambda^2p_t}\right) \tag{1a}$$

$$\Gamma(r) = \Gamma - C_\Gamma(T,r) + 2L_{at} - (r) + 40\log_{10}(r_0) - + P_r(r_0) \tag{1b}$$

Equation (1b) links the calibration term $C_\Gamma(T)$ $C_\Gamma(T,r)$ to the RCS Γ $\Gamma(r)$ of a target at a distance r_0. Γ is usually r. $\Gamma(r)$ is expressed in $dBsm$ units (decibels referenced to a square meter), L_{at} $L_{at}(r)$ is the atmospheric attenuation between the object and the radar in $dB\ km^{-1}$ dB, which can be calculated using a millimeter-wave attenuation model (for ex. (Liebe, 1989))and $P_r(r_0)$, $P_r(r)$ is the power received from the target in unreferenced dB $dB(AU)$ and $C_\Gamma(T,r)$ is the RCS calibration term in $dB(m^{-2}AU^{-1}) = 10\log_{10}(m^{-2}AU^{-1})$ units. The units in the RCS calibration term compensate the radar power units, guaranteeing the retrieval of physical RCS values. The explicit temperature and range dependency of the calibration term has the function of compensating gain changes in $P_r(r)$ introduced by temperature effects and variations in the IF loss with distance.

This principle can be used in an end-to-end calibration by installing a target with a known RCS Γ_0 at a known distance r_0 and sampling the power $P_r(r_0)$ reflected back to calculate $C_\Gamma(T,r)$. However, some additional considerations must be made to perform this retrieval.

In Eq. (1a) we state that the calibration value has a temperature dependency. This T dependency of the gain is reflected in Eq. (1b) as variations in the value of $P_r(r_0)$. and a range dependency. Experimental results indicate that $C_\Gamma(T)$ the temperature dependency of $C_\Gamma(T,r)$ can be approximated by a linear relationship, as shown in Eq. (2). Here n is the temperature dependency term in $dB\ {}^\circ C^{-1}$, T the internal radar temperature in $^\circ C$, and T_0 a reference temperature value in $^\circ C$ and C_Γ^0 a term we name the *calibration coefficient*, in dB. More details about this approximation the temperature correction can be found in Sect. 5.4(5.4).

$$C_\Gamma(T) = C_\Gamma^0 + n(T - T_0)$$

The range dependency of $C_\Gamma(T,r)$ is treated independently, by defining a IF loss correction function $f_{IF}(r)$, in dB units. This function is introduced to compensate for relative loss variations at different distances. The IF loss correction function is studied in Sect. 5.5.

Each $P_r(r_0)$ measurement is associated to a $C_\Gamma(T)$ value with From the afforementioned observations, we divide $C_\Gamma(T,r)$ in three components, shown in Eq. (4b). Then, with the temperature dependency n known, we use Eq. 2). This separation consists of a constant calibration coefficient C_Γ^0, in $dB(m^{-2}AU^{-1})$, and the two correction functions $n(T - T_0)$ and $f_{IF}(r)$.

$$C_\Gamma(T,r) = C_\Gamma^0 + n(T - T_0) + f_{IF}(r) \tag{2}$$

As $f_{IF}(r)$ corrects for relative variations in receiver loss with distance, we define $f_{IF}(r_0) = 0$ at the reference reflector position r_0. Using this and Eqs. (1b) and (2)to compute a , we get Eqs. (3a) and (3b).

$$C_\Gamma(T,r_0) = C_\Gamma^0 + n(T - T_0) \tag{3a}$$

$$C_\Gamma(T,r_0) = \Gamma_0 - 40log(r_0) - 2L_{at}(r_0) - P_r(r_0) \tag{3b}$$

Equation (3a) shows how the calibration term $C_\Gamma(T,r_0)$ at position r_0 is related to the calibration coefficient C_Γ^0 sample. After, n and and the temperature correction $n(T - T_0)$. Meanwhile, Eq. (3b) indicates how experimental $P_r(r_0)$ measurements can be associated with a $C_\Gamma(T,r_0)$ value, using in-situ information to calculate $2L_{at}(r_0)$. Then, using Eq. (3a), we can compute

C_Γ^0 are used to calculate $C_\Gamma(T)$ and then by substracting the temperature correction function $n(T-T_0)$. This temperature correction is derived independently in Sect. (5.4). Knowing C_Γ^0 and the temperature correction, $C_\Gamma(T,r)$ is calcualted by adding the IF correction function, independently retrieved in Sect. 5.5.

Once $C_\Gamma(T,r)$ is known, we can calculate the radar Equivalent Reflectivity calibration term $C_Z(T)$, in $dB$$C_Z(T,r)$, in $dB(mm^6m^{-5}AU^{-1})$, with Eq. (4a) (Yau and Rogers, 1996). This relationship assumes the radar has two identical parallel antennas with a Gaussianly shaped main lobe. Θ θ is the antenna beamwidth in radians, δr $m\delta r$ is the radar distance resolution in meters and $|K| = |(\epsilon_r - 1)/(\epsilon_r + 2)|$ is a parameter the dielectric factor. This factor is related to the complex dielectric constant relative complex permittivity ϵ_r of the scattering particles(for weather radar usually liquid water or ice), and can be calculated, for example, using the results of Meissner and Wentz (2004).

As with $C_\Gamma(T)$, $C_Z(T)$ $C_Z(T,r)$ enables the calculation of the Radar Equivalent Reflectivity term Z_e , in dBZ units of a distributed target located at a distance r, by using the relationship of Eq. (4b).

$$C_Z(T,r) = 10\log_{10}\left(\frac{8\ln(2)\lambda^4 10^{18}}{\theta^2\pi^6K^2\delta r}\right) + C_\Gamma(T,r) \tag{4a}$$

$$= Z_e - (r) = C_Z(T,r) + 2L_{at} - (r) + 20\log_{10}(r) - + P_r(r) \tag{4b}$$

3 Experimental setup

Two calibration campaigns, that lasted one month each, were performed in May-June of 2018 and March-April of 2019 at the SIRTA observatory, located in Palaiseau, France (Haeffelin et al., 2005). The observatory has a 500 meter long grass field in an area free of buildings, trees or other sources of clutter, well suited to install our calibration setup, shown in Fig. 1.

The instrument used for the calibration experiments is a BASTA-Mini. BASTA-Mini is a 95 GHz FMCW radar with scanning capabilities and two parallel Cassegrain antennas (Delanoë et al., 2016). The antennas are separated by 35 cm, and have a Fraunhofer far field distance of ≈ 50 m with a Gaussianly shaped main lobe (verified experimentally in Sect. 5.2). Transmitted power is fixed to 0.5 W, and is under constant monitoring using a diode with an uncertainty of ≈ 0.4 dB. The diode enable the monitoring of $L_t(T)$ variations, yet our experiments have shown that T is a better indicator to capture the variability of $C_\Gamma(T)$$C_\Gamma(T,r)$. This is likely because internal temperature changes affect both $L_r(T)$ $L_r(T,r)$ and $L_t(T)$ simultaneously, and therefore the information provided by the diode is not sufficient to capture the behavior of the whole system. The results of the temperature dependency study for our radar is shown in Sect. 5.4).

This radar also includes hardware to enable the tuning of the carrier wave frequency within a range of ≈ 1 GHz, centered at 95 GHz. During the experiments we fixed the BASTA-Mini base frequency at 95.64 GHz to avoid any interference with the other two W band radars operating in parallel at the same site.

Our reference targets are two Triangular Trihedral Reflectors (also known as Corner Reflectors) composed by three orthogonal triangular conducting plates. Trihedral targets have a large RCS for their size and a low angular variability of RCS around their boresight (Atlas, 2002; Doerry and Brock, 2009; Chandrasekar et al., 2015). One reflector has a size parameter of 10 cm,

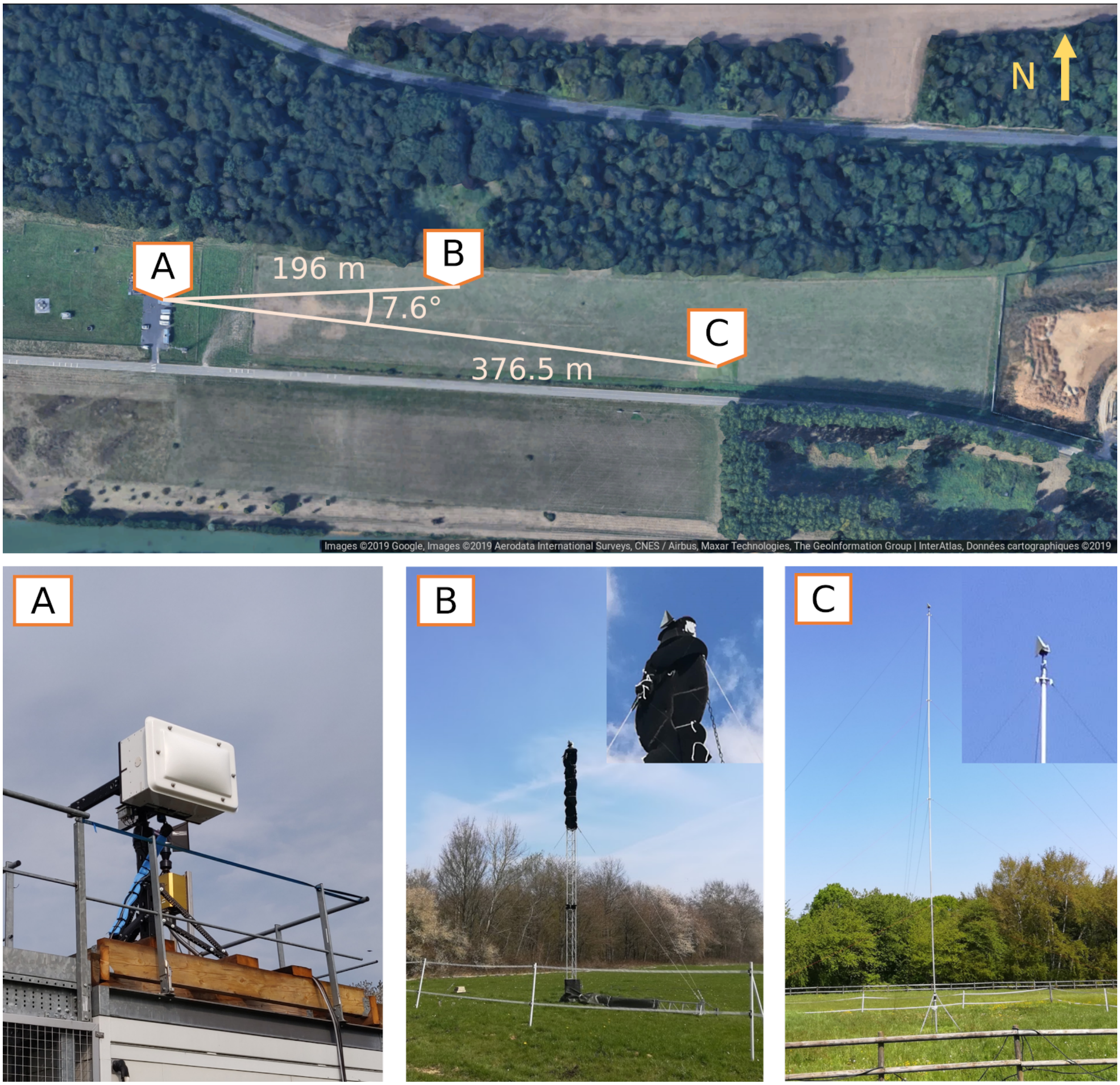

Figure 1. Experimental setup for 2018 and 2019 calibration experiments: (A) Scanning BASTA-Mini radar located in a reinforced platform 5 m above the ground. (B) 10 mast with a 10 cm triangular trihedral target mounted on a pan-tilt motor with an angular resolution and repeateability better than 0.1°. This mast has microwave absorbing material wrapped to it to reduce its RCS (clutter). The 10 m mast was only installed in the 2019 calibration campaign. (C) 20 m mast with a 20 cm triangular trihedral target. The target aiming is fixed relative to the mast. This mast is used in both 2018 and 2019 calibration campaigns. Angular separation between the masts is enough to sample both targets without mutual interference.

with a maximum RCS at our radar operation frequency of ~~42.63 m^2~~$16.30\ dBsm$. The other is 20 cm with a maximum RCS of ~~682.10 m^2~~ $28.34\ dBsm$ (Brooker, 2006). These targets were mounted on top of masts B and C in Fig. 1 respectively. Only mast C was used in the 2018 campaign, while both were used in 2019.

To align the system first we aim the radar towards the approximate position of the target. Second, we aim the target by slowly changing pan-tilt angles in the motor on mast B, or axially rotating the tube of mast C to maximize ~~power~~ the power $P_r(r_0)$ measured at the radar~~'s end~~. Third, radar aiming is tuned around target position until the maximum reflected power is found. Finally, we repeat the second step, after which we have the system ready to sample $P_r(r_0)$.

It must be mentioned that this procedure does not guarantee a perfect alignment. In fact, it is impossible to have every element perfectly adjusted because of limits in the positioner resolution or uncertainties introduced when installing each element. Sections 4 and 5.6 explain how we deal with these limitations.

4 Methodology

This section describes the procedure followed when performing calibration experiments using the setup described in Sect. 3. The methodology has the objective of quantifying and correcting when possible all sources of uncertainty to enable a reliable estimation of the calibration terms ~~$C_\Gamma(T)$ and $C_Z(T)$~~$C_\Gamma(T,r)$ and $C_Z(T,r)$.

A challenge we found when using targets mounted on masts to estimate ~~C_Γ~~ $C_\Gamma(T,r)$ is that the value of ~~Γ~~ the target RCS Γ_0 may vary depending on how components are aligned. Our studies have shown that for the feasible alignment accuracy we can get when installing our setup, this effect is in the order of tenths of dB, and therefore not negligible. Additionally, we concluded that if we ~~left~~ leave this uncertainty source uncorrected, we would introduce a bias in the calibration result (see Sect. 5.6).

The flow chart of Fig. 2 illustrates the calibration procedure. To quantify the bias introduced by alignment uncertainty we decided to divide each calibration exprediment in N iterations. Each iteration consists on a system realignment, followed by sampling of the target signal $P_r(r_0)$ for at least one hour. Then, we select the data from the contiguous hour with the lowest variability as the iteration result.

The period chosen to perform the sampling is important, because it will have an incidence on how stable is the calibration value. To minimize uncertainty it is recommended to perform calibration iterations when the atmosphere is clear, there is no rain and wind speed is under 1 ms^{-1}. However, these requirements may change depending on how robust is each setup to atmospheric conditions.

FMCW radars have a discrete distance resolution. Consequently, power measurements vs distance ~~will be~~ are resolved in finite discrete points, usually named gates. Because of this resolution limitation, power received from a point target is ~~split~~ spread between the gates closer to its position (Doviak and Zrnić, 2006). This phenomena is known as spectral leakage. To reduce leakage BASTA-Mini uses a Hann time window (Richardson, 1978; Delanoë et al., 2016).

To correctly asses the total reflected power we set the radar resolution to 12.5 meters (chirp bandwidth of 12 MHz), and its integration time to 0.5 seconds. This resolution is high enough to accurately identify the ~~peak of power coming from the target~~

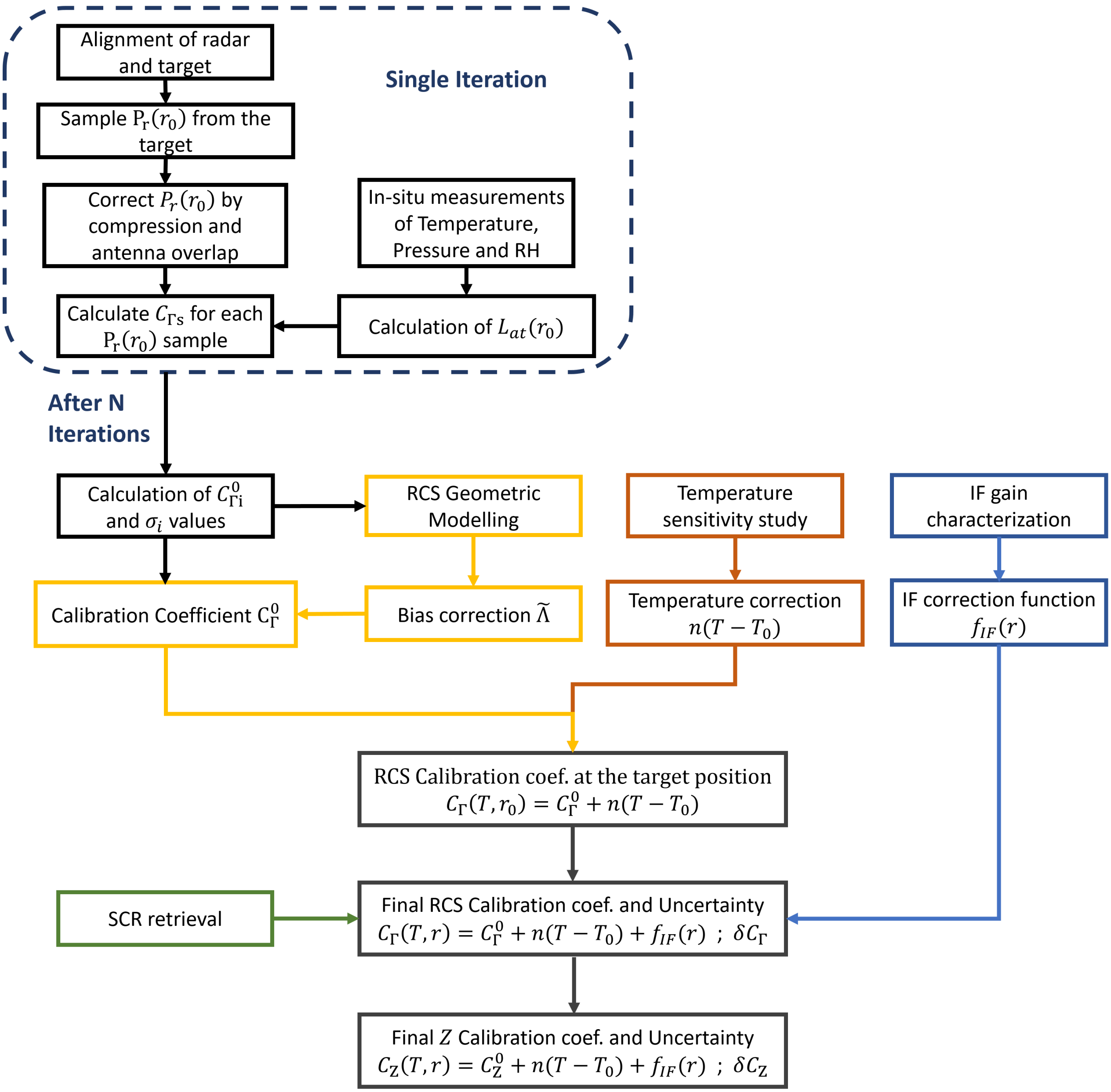

Figure 2. Summary of a complete calibration process. Each calibration requires repetition of system realignment and sampling steps, called iterations. During each iteration we continously sample the power reflected from the reference target position for one hour (power corrections in Sect. 5.1). The retrieval of N iterations enable the estimation of the system bias due to misalignments in the setup (Sect. 5.6). Temperature dependency is retrieved in an independent experiment (Sect. 5.4). Uncertainty introduced by clutter signals at the target location is also included in the total uncertainty budget (Sect. 5.3).

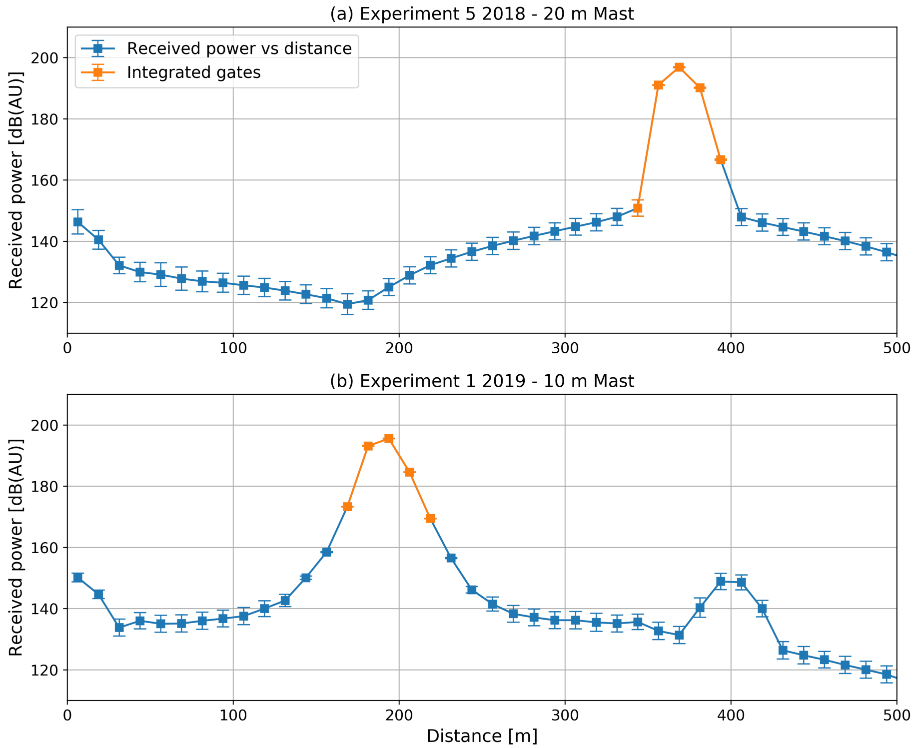

Figure 3. Mean profiles of received power for Experiment 5 in 2018 using the 20 m mast (a) and Experiment 1 in 2019 using the 10 m mast (b). Standard deviation at each gate is indicated with an errorbar. The gates integrated to calculate the reference reflector backscattered power $P_r(r_0)$ are marked in orange. The secondary peak of figure (b), around 400 meters, corresponds to reflections on trees behind the 10 m mast.

reference reflector signal while avoiding the introduction of additional clutter from the trees located ~~further~~behind the mast (see Fig. (3)).

To calculate $P_r(r_0)$ we add five gates: the target's gate plus two before and two after the target's position. Adding more contiguous gates increase the power value by less than $0.01\ dB$, thus we conclude that these five gates concentrate almost all the power reflected back from the target.

Then $P_r(r_0)$ is corrected considering compression effects and antenna overlap losses (Sects. 5.1 and 5.2). For each corrected $P_r(r_0)$ sample we proceed to calculate a single ~~C_Γ value~~ C_Γ^0 value with Eq. (3a) and the temperature correction function. This single sample is defined as $C_{\Gamma s}^0$ to differentiate it from the final calibration coefficient C_Γ^0 of Eq. (2). Atmospheric attenuation ~~L_{at}~~ $L_{at}(r_0)$ is calculated using in-situ atmospheric observations and the model published by Liebe (1989).

The target effective RCS Γ_0 is calculated using a theoretical RCS model, considering the beam incidence angle on the target but neglecting possible alignment errors. It is worth noting that echo chamber measurements have shown that real targets RCS can be deviated from the theoretical value depending on the manufacturing precision. Our corner reflectors have an angular manufacturing precision better than $0.1°$, therefore RCS deviations could be on the order of 1-2 $dBsm$ (Garthwaite et al., 2015) . However, since at the writing time we do not have an experimental characterization for our targets, we rely on the theoretical model. This is not a major issue because, once an experimental characterization of the target becomes available, it can be used to correct any calibration bias by rectifying the value of Γ_0 used in the calculations.

We performed one calibration experiment with 6 iterations during the 2018 campaign using the 20 m mast. In the 2019 campaign we did two experiments: one with 10 iterations using the 10 m mast and another with 2 iterations on the 20 m mast (Fig. 1). During these campaigns we also retrieved

The retrieval of the temperature dependency coefficients coefficient n and the reference temperature T_0 experimentally (see is done simultaneously with the calibration coefficient experiment, by extending the sampling period beyond one hour when using the 20 m mast. This is done to capture the temperature effect in the variability of $C^0_{\Gamma s}$, by capturing a larger part of the temperature daily cycle. The results of this experiment can be seen in Sect. 5.4). Temperature dependency coefficients enable the calculation of a C^0_Γ value for each sampled C_Γ. . Likewise, the retrieval of the IF correction function $f_{IF}(r)$ is an independent experiment based on sampling noise with the radar to get the IF amplification curve of the receiver. The details of this experiment are in Sect. 5.5

From each iteration we get a distribution of resulting C^0_Γ $C^0_{\Gamma s}$ values with a small spread introduced by second order effects. The average value of each iteration i is named $C^0_{\Gamma i}$, and its corresponding standard deviation is named σ_i σ_i. With this information we proceed to calculate the bias corrected calibration coefficient C^0_Γ, by using Eq. (5). $\tilde{\Lambda}$ is the bias correction term. The method used to calculate $\tilde{\Lambda}$ relies on simulating the probability distribution of Γ_0 for a given set of uncertainties in the setup parameters. More detail can be found in Section 5.6 and Section S3 of the supplementary material.

$$C^0_\Gamma = \frac{1}{N}\sum_{i=1}^{N} C^0_{\Gamma i} - \tilde{\Lambda} \tag{5}$$

Equation (6) shows the uncertainty budget linked to the afforementioned estimator of C^0_Γ. δC_Γ and δC_Z are the resulting uncertainties of $C_\Gamma(T)$ and $C_Z(T)$ respectively. The terms σ^2_T/N Equations (6a) and (6b) show the uncertainties δC_Γ and δC_Z associated with the estimation of $C_\Gamma(T,r)$ and $C_Z(T,r)$ respectively.

σ_T is the uncertainty term associated with the temperature correction function $n(T-T_0)$.

σ_{IF} is the uncertainty term associated with the IF loss correction function $f_{IF}(r)$.

The term $\sum\sigma^2_i$ come from propagating the uncertainty when averaging comes from the averaging operation in the estimation of $C^0_{\Gamma i}$ (Eq. 5). Since the $C^0_{\Gamma i}$ terms . σ_Λ are corrected using the temperature correction function, the uncertainty of the later must be propagated as well, hence the term σ^2_T/N appears.

σ_Λ the uncertainty of the bias correction and σ_{SCR} calculation. It is calculated from the standard deviation σ_i. This procedure is explained in Section S3 of the supplementary material.

σ_{SCR} is the uncertainty introduced by clutter. Clutter is the presence of unwanted echoes which affect our reading of $P_r(r_0)$, coming from reflections on other objects in the environment. The method to quantify the uncertainty σ_{SCR} uses a parameter named Signal to Clutter Ratio (SCR), explained in detail in Sect. 5.3. Finally, the additional σ_T term appears when calculating $C_\Gamma(T)$ or $C_Z(T)$ with Eq. (2) during the radar operation.

$$\delta C_\Gamma = \sqrt{\frac{1}{N^2}\sum_{i=1}^{N}\sigma_i^2 + \frac{\sigma_T^2}{N} + \sigma_T^2 + \sigma_{SCR}^2 + \sigma_\Lambda^2}$$

$$\delta C_Z = \delta C_\Gamma$$

We additionally observe that RCS and Z_e calibration uncertainty is the same. This comes from the following observations: Wavelength variations coming from the chirp and distance resolution uncertainties contribute less than $0.01\ dB$ to the uncertainty budget. Uncertainty introduced by assuming a Gaussian beam shape is neglected, because the RMSE between the Gaussian fit and the manufacturer's σ_{Γ_0} is the uncertainty of the reference target RCS. In this work we use a theoretical model to calculate the target effective RCS. The inclusion of an experimental characterization of the antennas is less than $0.01\ dB$ within the HPFW region (see Sect. 5.2). target RCS can improve the estimation of C_Γ^0 and δC_Γ. Because we do not have this information available at the writing date, we neglect this term.

It is worth noting that ϵ for water at the W band can vary significantly with temperature. Using the results published by Meissner and Wentz (2004), we calculate that $|K|$ varies from ≈ 0.83 σ_K is the uncertainty in the estimation of the backscattering particles dielectric factor. Because our objective is to calculate the calibration term of the radar, we reference this value to ≈ 0.92 between -5 and 25 ^{o}C. This variability translates as a change of -0.5 to $+0.5$ dB in C_Z, respectively. For this study we 'll use the reference value $|K| = 0.86$, corresponding to pure water at 5 ^{o}C, however if lower uncertainty in Z_e retrievals is necessary this value should be corrected by using in-situ or remote sensing retrievals of cloud temperature.The same applies to $^\circ C$ and neglect the δ_K uncertainty term. However, the value of K and its uncertainty σ_K must be considered when performing radar retrievals (e.g. Sassen (1987); Liebe et al. (1989); Gaussiat et al. (2003)).

σ_A is the uncertainty introduced in the estimation of L_{at}. Thus, uncertainty in the calibration is not the same as the uncertainty in the retrievals but rather its lower boundθ and from parallax errors and deviations from a Gaussian beam shape (Sekelsky and Clothiaux, 2002). For this work we make the assumption of parallel antennas with a Gaussian beam shape, thus we neglect this term. This problem is dissused more in depth in Section 5.2.

Since both σ_K and σ_A are neglected, we get $\delta C_\Gamma \approx \delta C_Z$.

$$\delta C_\Gamma = \sqrt{\frac{1}{N^2}\sum_{i=1}^{N}\sigma_i^2 + \frac{\sigma_T^2}{N} + \sigma_{IF}^2 + \sigma_T^2 + \sigma_{SCR}^2 + \sigma_\Lambda^2 + \sigma_{\Gamma_0}^2} \tag{6a}$$

$$\delta C_Z = \sqrt{\delta C_\Gamma^2 + \sigma_K^2 + \sigma_A^2} \tag{6b}$$

5 Sources of uncertainty and bias in Absolute Calibration with corner reflectors

In this section we identify and quantify the uncertainty and bias introduced by several terms in Eq. (1b). Following the recommendations in the work of Chandrasekar et al. (2015), we ~~considered~~ study the impact of receiver saturation, signal to clutter ratio, antenna lobe shape and ~~overlap, and environmental conditions~~antenna overlap. Additionally, we ~~considered~~ consider the impact of temperature fluctuations inside the radar box~~and~~ , loss changes with distance due to uneven amplification at the receiver IF and the effects of imperfect alignment of the reference target.

5.1 Receiver compression

It is advisable to design calibration experiments which avoid the appearence of compression effects. If this is not possible, compression must be considered in the data treatment so that the retrieved calibration remains valid in the receiver's linear regime, where it usually operates during cloud sampling (Scolnik, 2000).

For studying how these effects could affect our calibration, we retreived the radar's receiver power transfer curve. Receiver characterization was done by removing the radar's antennas and connecting the emitter's end to the receiver's input, with two attenuators in between. The first was a $40\ dB$ fixed attenuator, while the second was a tunable attenuator covering the range between 50 and $1\ dB$ of losses. The adjustable attenuator enabled the retrieval of the power transfer curve by varying the attenuation and sampling the power at the receiver's end (digital processing included). Our retrieved power transfer curve is shown in Fig. 4 (a).

Compression effects must be considered in calibration, or a bias will be introduced. In consequence, we include compression correction in every sample of reflected power, which consists on projecting their value to the ideal linear response using the power transfer curve.

For example, the power received from the 20 cm target on the 20 m mast returned was $198.7\ dB$ in average, before corrections. The power transfer curve shows that at this power values we have a loss caused by compression of $\approx 0.3\ dB$. After correcting each power sample by compression with the power transfer curve, we obtain a corrected power average value of $199.1\ dB$. On the other hand, for the 10 cm target on the 10 m mast the average power value before corrections is 197.8. ~~Because~~ As this value is lower than what is obtained the 20 m mast, the associated compression effect is also smaller, of $\approx 0.2\ dB$. After aplying this correction to each power sample we end with a new corrected power average of $198.0\ dB$.

5.2 Antenna Properties

~~We took advantage of our experimental setup and the scanning capabilities of the radar to check if the radar antennas were properly aligned and if their beam width matched the specifications provided by the manufacturer. This was done by using the target on the 20 m mast. Results are shown in Fig. 4 (b).~~

Manufacturer specifications indicate that antenna beamwidth should be of $0.8°$. However, data from an experimental characterization done by the same manufacturer in an anechoic chamber indicate that antenna beam shape is better approximated by a Gaussian with a Half Power Beam Width (HPBW) of $\theta \approx 0.88^o$. The total gain difference between the experimental curve and the Gaussian approximation of $\approx 0.0003\ dB$ in the HPBW region.

The beamwidth of 0.88° is consistent with what we observe in our scanning results. This indicates that we can assume the antennas to be parallel , allowing the introduction of Therefore, we conclude that the contribution to uncertainty introduced by assuming a Gaussian beam shape is negligible. The Antenna beam shape and Gaussian curve are shown in Fig. 4 (b).

Another source of bias introduced by the antennas is the parallax error. Antenna parallax errors introduce a range dependent bias, determined by the antenna beamwidth and the relative angles of deviation between the antennas boresight. This bias is usually larger in the first few hundred meters closest to the radar. For example, for a deviation of half the antenna beamwidth, losses would be on the order of 10 dB and would vary significantly over the first hundreds of meters, decreasing with distance to about 1 dB at a approximately 4 kilometers (Sekelsky and Clothiaux, 2002).

To study this effect we took advantage of our experimental setup and the scanning capabilities of the radar, to check if the radar antennas were properly aligned. This was done by using the target on the 20 m mast. Results are shown in Fig. 4 (b). After analyzing the results we observed that the aiming uncertainty is in the same order of magnitude of the antennas beamwidth. Since the correction of the parallax error requires a very precise measurement of antenna alignment, we conclude that it is not possible to correct for antenna deviations directly with this information.

However, the relativelly small difference of 0.5 dB in the estimation of C_Γ^0 during the calibration experiments of 2019, obtained using two masts in the most sensitive distance range (placed at 196 and 376.5 meters of distance respectively), indicate that antennas are unlikely to have a deviation comparable to their beamwidth (calibration results in Sect. 6).

Therefore, for the present version of this calibration methodology we assume that both antennas are parallel and that they have a Gaussian beam lobe. Once a reliable method for antenna pattern retrieval is developed for W band radars, it can be directly incorporated into the calibration term by adding an additional correction function $f_A(r)$ to Eq. (2). The uncertainty in this alignment estimation can also be included in the uncertainty budget, with the term σ_A of Eq. (6b).

Even if antennas are parallel, it is necessary to include a correction for losses the loss $L_o(r)$ caused by incomplete antenna overlap. The correction, shown in Eq. (7), accounts for the loss of power that would be received from a point target compared to a monostatic system (Sekelsky and Clothiaux, 2002). This loss occurs because a point target cannot be in the center of two non-concentric parallel antenna beams.

$$L_o(r) = \exp\left(\frac{2\arctan(\frac{d}{2R})^2}{0.3606\ \theta^2} \frac{2\arctan(\frac{d}{2r})^2}{0.3606\ \theta^2} \right) \tag{7}$$

Equation (7) assumes that the radar has two identical, parallel antennas with gaussian beam shapeGaussian beam lobes. Their main axis is separated by a distance d, and the point target is located at a distance $R r$, facing the geometrical center of the radar, where the gain is maximum. For the BASTA-Mini $d = 35\ cm$. This introduces a loss of $0.08\ dB$ for the target at 196 $r_0 = 196$ meters of distance, and of $0.02\ dB$ for the target at 376.5 $r_0 = 376.5$ meters.

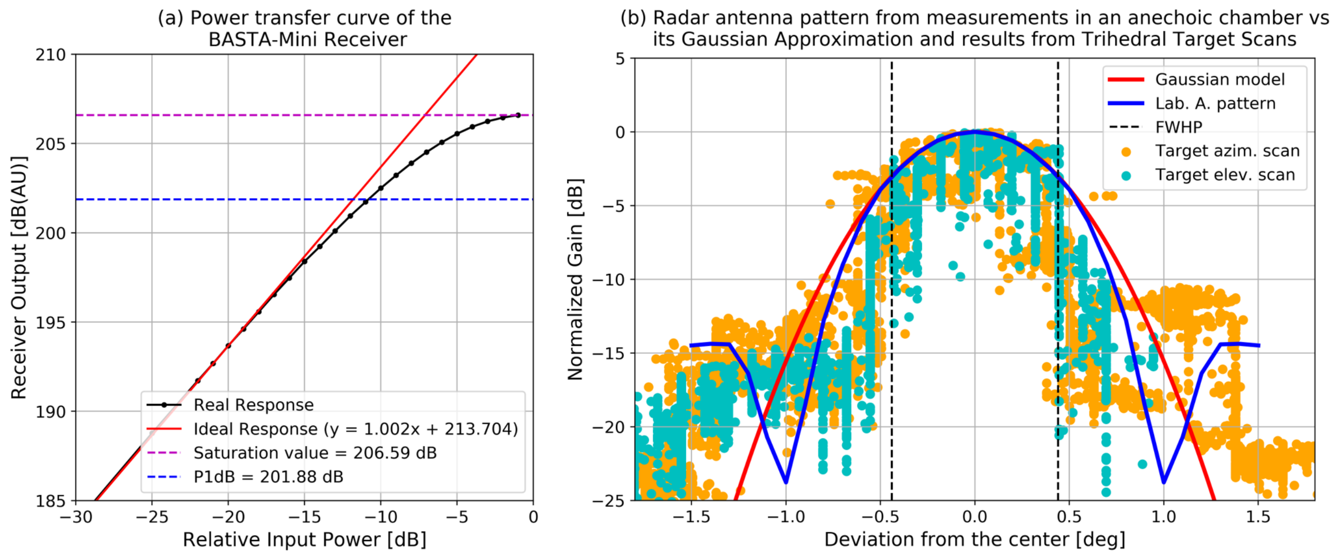

Figure 4. (a) Power transfer curve of the BASTA-Mini receiver. Input power is relative to the minimum attenuation value of the curve characterization experiment. All our signal retrievals from the target are slightly under 200 [~~dB~~$dB(AU)$] line, thus the correction required due to compression effects is small ($< 0.3\ dB$). (b) Normalized antenna pattern of the BASTA-Mini antennas. We can observe that the Gaussian fit with a beamwidth of $\theta = 0.88°$ is very close to the antenna gain curve measured at the manufacturer's laboratories~~and~~. This figure also shows the results from ~~our~~ mast scans ~~, specially within~~ around the ~~HPBW area (3 dB loss)~~target to compare with the theoretical curves. ~~This fit has a beamwidth~~ To enable the comparison with the laboratory antenna pattern we assume that the gain of ~~$\theta = 0.88°$~~both antennas is identical. Then, the received power in $dB(AU)$ is normalized with respect to the maximum measured value and divided by 2, to represent the gain of a single antenna.

5.3 Signal to Clutter Ratio

The power sampled from our reference reflector is an addition of the power from the target (signal) and unwanted reflections on other elements in the environment, such as the ground or the mast (clutter). We observed that this clutter dominates above the radar noise, and thus becomes the main source of interference in our calibration signal.

To quantify the impact of clutter we use the Signal to Clutter Ratio (SCR) parameter. It is calculated as the ratio of total power received from the target to power received from clutter under the same configuration, but with the reference reflector removed. SCR enables the uncertainty σ_{SCR} introduced by clutter in the sampled $P_r(r_0)$ values to be computed (Chandrasekar et al., 2015).

Clutter power is sampled and corrected ~~with the same corrections used in~~ following the same methodology used for reflector $P_r(r_0)$ retrievals, but in an scanning pattern mode to capture clutter around the ~~whole target~~ mast area. Figure 5 shows our results from scanning around the 10 and 20 m masts with targets removed.

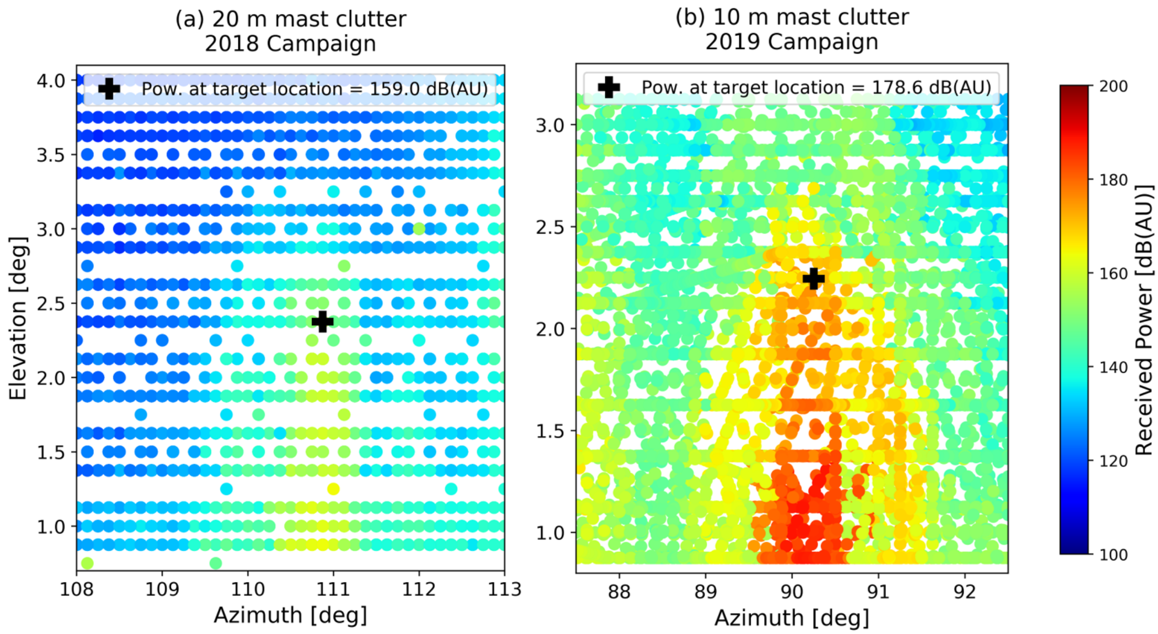

Figure 5. Clutter retrieval from the 10 m (a) and 20 m mast (b) respectively. Masts are scanned without the reflectors to measure the clutter signal. The nominal target position is marked with a black cross.

We observe that the 10 m mast is more reflective than the 20 m one. This may be caused by its smaller height (more ground clutter) and its larger geometrical cross-section. We can also see that the signal in the 10 m is stronger where absorbing material is not present (below $\approx 1.5°$ of elevation). In both cases we ~~didn't~~ did not detect any signal from the nearby trees close to the target's position.

To calculate SCR we compare the average power received from each target during the calibration experiments with the maximum clutter power observed in a region of 0.125^o around the target's coordinates, vertically and horizontally. The value is taken from the radar's positioner resolution.

The average reflected power from 10 cm target on the 10 m mast is 198.0 dB. This provides an SCR value of 19.4 dB, which implies a σ_{SCR} uncertainty value of ≈ 0.93 dB. From the 20 cm target on the 20 m mast, the average reflected power is 199.1 dB. Its SCR equals 40.1 dB, which is translated as an uncertainty contribution of $\sigma_{SCR} \approx 0.09\ dB$. From the results we see that even if target alignment is better with the 10 m mast, calibration results may not get less uncertain because the motor used for target alignment acts as a big source of clutter.

5.4 Temperature correction

BASTA-Mini has a regulation system to control temperature fluctuations inside the radar box. However, since the radar is based on solid state components, even small temperature fluctuations may impact the performance of the transmitter and receiver, and therefore affect the calibration stability. To account for this effect we introduced a temperature dependency in the calibration term, shown in Eq. (2).

During the experiments we verified the need of this correction by observing that the retrieved calibration term $C_\Gamma(T)$ $C_\Gamma(T,r_0)$ has a consistent change depending on the time of the day, and that this change is strongly correlated to the temperature inside the radar.

Figure 6 (a), (b) and (c) show the results of a representative experiment done in the 2018 campaign. Here we left the radar sampling the target signal for several hours, to observe the variability of $C_\Gamma(T)$ $C_\Gamma(T,r_0)$ during the day. (a) shows the raw result in the RCS calibration term $C_\Gamma(T)$$C_\Gamma(T,r_0)$. There is a spread of almost 1 dB between the maximum and minimum values during the whole timeseries. (b) is a fourier transform of this raw timeseries. Here we can see that most of the variability happens in the timescale of hours. (c) presents the timeseries of (a), but in a daily cycle perspective. Here we plot hourly means of the deviation of $C_\Gamma(T)$ $C_\Gamma(T,r_0)$ with respect to the total average, with its hourly standard deviation as errorbars. We also superimposed atmospheric attenuation and the radar amplifier temperature to show that the first has a much smaller impact in calibration variability compared to the second.

Figure (d) shows the raw results of plotting variations in $C_\Gamma(T,r_0)$ to temperature changes around $T_0 = 26.5\ ^oC$. These variations are calculated independently for each iteration, by substracting the constant term of the linear fit of $C_\Gamma(T,r_0)$ with respect to temperature. This operation removes the effect introduced by differences in alignment between different iterations. The reference T_0 value is chosen because it is approximately the average internal temperature when considering all the experiments.

To retrieve the temperature dependency we use a linear regression over the results from all the experiments done in 2018 and 2019, as shown in Fig. 6 (d)7. In this case the data used was not limited to one hour, to maximize the range of temperatures covered. The regression shows that the variability in the calibration term has an almost linear relationship with internal radar temperature, in the dB scale, and it is the same for both campaigns.

Study of calibration variability. Samples from iteration 5, 2018 Calibration Campaign. (a) Time series of the RCS Calibration term retrieval. (b) Fourier transform of the RCS Calibration term after substracting the mean value. (c) Daily Cycle of calibration variability, amplifier temperature and two-way attenuation. Attenuation errorbars are too small to be seen with this scale. (d) Regression of the relative changes in the $C_\Gamma(T)$ versus amplifier temperature, calculated using all samples from 2018 and 2019 campaigns.

The analysis enable us to obtain estimations for the terms of the temperature dependency part This analysis allows us to estimate the value $n = 0.093\ dB\,^\circ C^{-1}$ for the temperature correction function of Eq. (2). We obtain the value $n = 0.093$ $dB\,^\circ C^{-1}$, with $T_0 = 26.5\ ^oC$. This reference T_0 is chosen because it is approximately the average internal temperature when considering all the experiments.

When calculating To estimate the uncertainty of the temperature correction function we calculate the RMSE between the linear regression and all data, we get obtaining the value $\sigma_T = 0.13\ dB$. This remaining error may be introduced by the combination of random noise

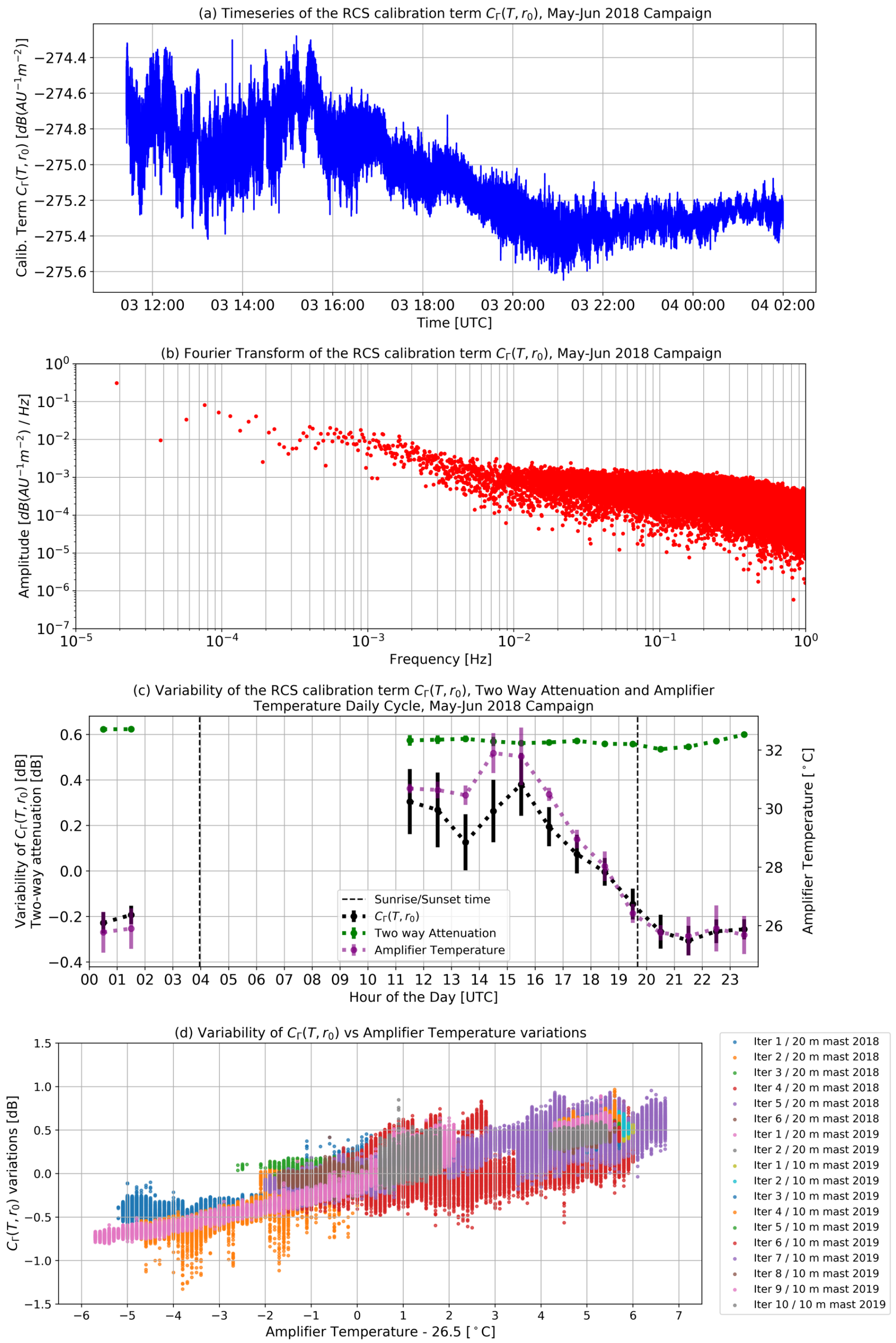

Figure 6. Calibration variability study. Samples from iteration 5, 2018 Calibration Campaign. (a) Time series of the RCS Calibration term retrieval. (b) Fourier transform of the RCS Calibration term after substracting the mean value. (c) Calibration variability daily cycle, amplifier temperature and two-way attenuation. Attenuation errorbars are too small to be seen with this scale. (d) Relative changes in $C_\Gamma(T, r_0)$ versus amplifier temperature, plotted using all samples from 2018 and 2019 campaigns.

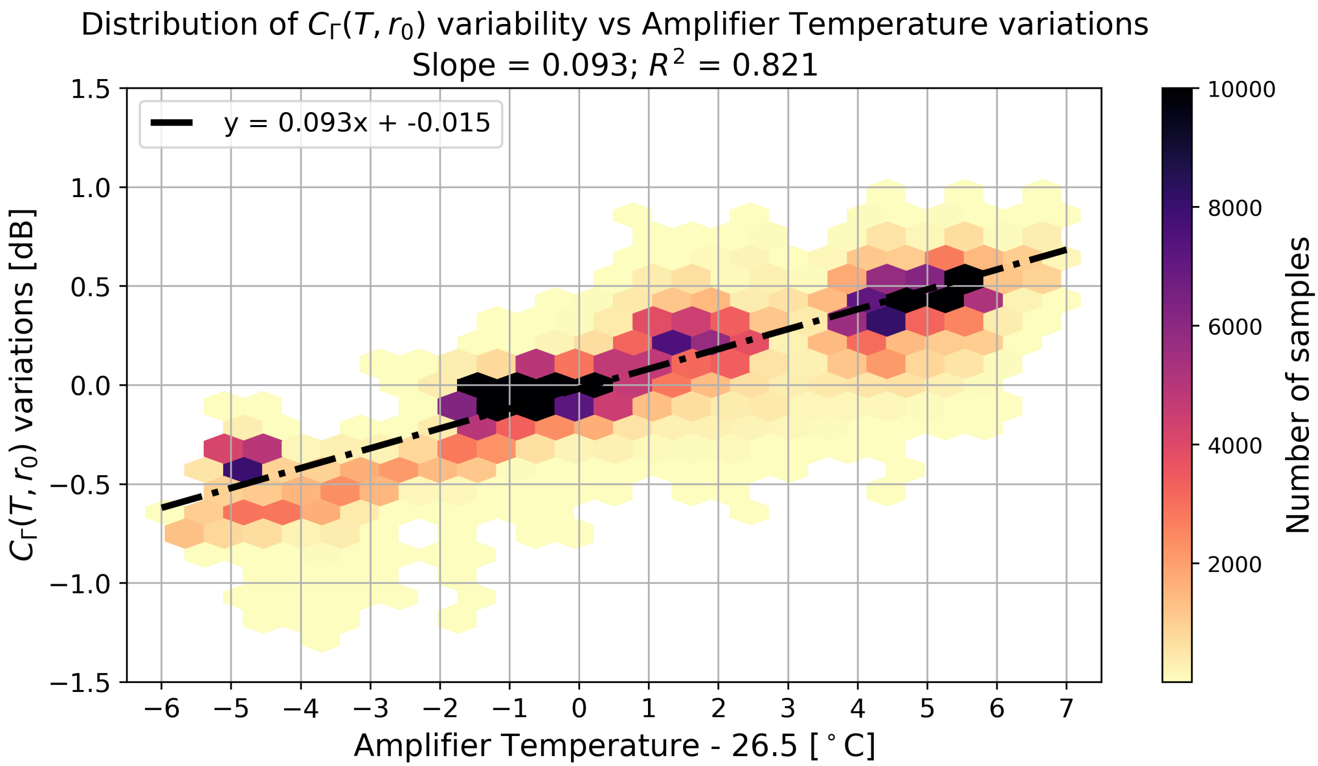

Figure 7. 2D histogram of the relative changes in $C_\Gamma(T, r_0)$ with respect to changes in the amplifier temperature and its linear least squares fit. The histogram is plotted using all $C_\Gamma(T, r_0)$ samples from 2018 and 2019 calibration campaigns.

5.5 IF loss correction function $f_{IF}(r)$

FMCW radars rely on estimating the beat frequency of the received signal to estimate the distance of an object. This signal may suffer uneven amplification depending on its frequency, because of a frequency dependent gain function in the amplifiers of the IF chain of the radar. Since there is a direct relationship between the beat frequency and the target distance r, this dependency on the beat frequency introduces a gain variability with respect to the target distance r. As introduced in Sect. 2, this distance dependency is compensated in the calibration term with a IF correction function $f_{IF}(r)$

To retrieve the $f_{IF}(r)$ we turn off the radar emitter and sample the environmental noise with the radar operating in its calibration mode (12.5 meters distance resolution and 0.5 seconds integration time). The power $P_r(r)$ measured by the radar under these conditions corresponds to the environmental noise power density N_0 reduced by the radar total receiver loss

$\log_{10}(L_r(T,r))$, as indicated in Eq. (8) (Pozar, 2009). Noise power density is assumed to be constant in the 12 MHz bandwith of this operation mode.

$$P_r(r) = N_0 - 10\log(L_r(T,r)) \tag{8}$$

After retrieving a significant amount of noise samples we calculate the average value of the difference $P_r(r_0) - P_r(r)$ for each distance r, to remove the effect of the unknown noise power density. By using Eqs. (1a) and (2), we get that the difference $P_r(r_0) - P_r(r)$ is equivalent to the difference between $C_\Gamma(T,r)$ and other non identified second order sources, so we included it in the final uncertainty budget of the calibration$C_\Gamma(T,r_0)$, and therefore it is equivalent to the IF correction function $f_{IF}(r)$. The temperature effect in gain is removed because both $P_r(r_0)$ and $P_r(r)$ are sampled simultaneously, and therefore under the same temperature conditions.

For this experiment only, $P_r(r_0)$ corresponds to the power measured at the gate closer to the reference target position, without integrating other gates. This is done because there is no significant leakage and, as results of Fig. (8) show, $L_r(T,r)$ changes are negligible in the five gates used for integration.

$$\overline{P_r(r_0) - P_r(r)} = -10\log(L_r(T,r_0)) + 10\log(L_r(T,r)) = -C_\Gamma(T,r_0) + C_\Gamma(T,r) = f_{IF}(r) \tag{9}$$

It must be noted that the IF correction function retrieved in this way is only valid for the operation mode used during its retrieval. If the radar operates in another mode, this function should be retrieved again, or a broader characterization of the IF loss must be used.

Figure 8 shows the results of the IF correction function retrieval referenced to $P_r(r_0)$ at $r_0 = 376.5$ meters (corresponding to the 20 m mast experiment setup). We can observe that all functions retrieved in 2019 are in close agreement, without significant variations between different dates or time of the day chosen for the plots. The 2018 function is different because hardware was modified between both calibration campaigns. Additionally, in 2018 the emitter was not turned off to perform the noise sampling. Rather, we resorted to use a sampling period with clear sky conditions to respect the assumption of Eq. (8).

A sixth degree polynomial is used to fit $f_{IF}(r)$ between 150 and 6000 meters, which is the maximum valid range for the 12.5 meter resolution mode (Delanoë et al., 2016). For both 2018 and all 2019 curves, the fit has a RMSE $< 0.03\ dB$. Furthermore, the standard deviation between results from the four periods of 2019 has a maximum value of $0.04\ dB$ for any gate. Both results indicate that the uncertainty introduced by the IF correction function is $\lesssim 0.04\ dB$. Finally, the IF correction function retrieved for the 10 m mast setup in 2019 (with $r_0 = 196$ meters) is almost identical to the 20 m mast results. These functions are presented in Sect. 6.

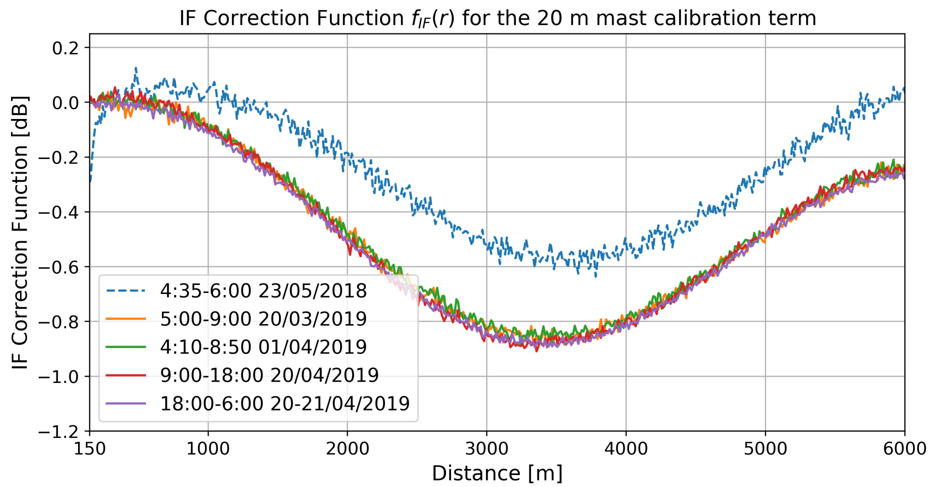

Figure 8. Data used for the IF correction function calculation, retrieved for different periods of the 2018 and 2019 calibration campaigns. 2018 curve is different from the 2019 results because hardware was modified between the campaigns. Time indicated in the label is in UTC.

5.6 Misalignment Bias

The retrieval of ~~$C_{\Gamma}(T)$~~ $C_{\Gamma}(T,r_0)$ using Eq. (~~1b~~3b) requires a precise knowledge of the reference target effective RCS Γ_0. Each $dBsm$ of difference between the theoretical value used in calculations and the effective target RCS will introduce a bias of the same magnitude in the estimation of the calibration ~~coefficients~~ coefficient C_{Γ}^{0}, and thus in ~~$C_{\Gamma}(T)$~~$C_{\Gamma}(T,r_0)$.

The effective reflector RCS is the actual physical value that would be measured by a perfectly calibrated radar. It is different from the target intrinsic RCS which only depends on its physical properties. Effective RCS changes when the experimental setup is modified. For example, if the point target is not exactly in the beam center, antenna gain ~~won't~~ will not be maximum and therefore the effective RCS will decrease compared to the intrinsic value. Effective RCS also changes when the incidence angle of the radar beam is modified. This latter effect may increase or decrease effective RCS depending on the original situation.

A common approach in these type of experiments is to set Γ_0 to be the maximum theoretical RCS of the target, assuming misalignment will cause a negligible deviation from this value. This procedure can be refined for cases where the system default configuration does not have the target boresight aligned with the radar position. In these cases, effective RCS can be calculated using equations derived from geometrical optics (more complex optical calculations may be necessary for other wavelengths or target sizes). For example, we use the equations published by Doerry and Brock (2009) when calculating the effective RCS of our Triangular Trihedral target on the 20 m mast.

Unfortunately, this approach does not correct the impact of alignment uncertainties. We observed that random errors in the element positioning will statistically impact the effective Γ_0 in a single direction. Thus, simply taking the average of many target sampling iterations would result in a biased estimation of the calibration.

With the objective of quantifying the impact of alignment uncertainties we developed a geometrical simulator of effective RCS. This simulator receives as input the position of each element in the setup and calculates the effective RCS considering the beam incidence angle and antenna gain variations when the target is not in the center of the beam. The degrees of freedom included in the simulator are shown in Fig. 9 (a). It enables the modification of the radar aiming angles, the mast dimensions and the positioning and orientation of the target. The equations used in the simulator can be found in the article support material.

We now use the simulator to study how uncertainty in alignment can affect the value of Γ_0. For this, we model an example experiment based on the 20 m mast setup. In this model we separate input variables between known and uncertain. Known terms can be fixed or measured very precisely in the field experiment, hence they are set as fixed values. Meanwhile, uncertain terms represent the parameters that cannot be fixed or measured very precisely, and for that reason are better expressed as probability distributions (terms defined in Fig. 9 (a)).

- Known terms:
  - $x_r = 376.5$ m
  - $h_r = 5.3$ m
  - $\rho = 20$ m
  - $\alpha = 48°$
  - Target Size = 20 cm
- Variables with uncertainty:
  - $\theta_r = \mathcal{N}(\theta_r^*, \sigma_{\theta_r}^2)$
  - $\phi_r = \mathcal{N}(\phi_r^*, \sigma_{\phi_r}^2)$
  - $\theta = \mathcal{N}(0, \sigma_\theta^2)$
  - $\phi = \mathcal{U}([0°, 360°))$
  - $\tau = \mathcal{N}(\tau^*, \sigma_\tau^2)$

In the uncertain variables, $\theta_r^* = 87.82°$, $\phi_r^* = 0°$ and $\tau^* = 0°$ represent the nominal alignment angles, which are the values expected under an ideal field experiment where the radar aims directly to the target and the mast is perfectly vertical. To these nominal values we associate a distribution shape and the uncertainty set $\sigma_{\theta_r} = 0.075°, \sigma_{\phi_r} = 0.075°, \sigma_\theta = 1.5°, \sigma_\tau = 5°$. Each term, known and uncertain, is estimated from observations done during the experimental field work.

With these input parameters we sample the Γ_0 distribution that would arise after a large amount of experimental iterations. Figure 9 (b) shows the results from this sampling. The black dashed line shows the effective RCS under our experimental configuration, when each element is in its nominal position. We can see that this effect cannot be neglected in our case, since its value is $0.8\ dB$ lower than the maximum theoretical RCS.

However, this single correction does not suffice. The results of the model show that the addition of uncertainty into the process induces another bias of $\approx 0.3\ dB$, in average. Since this is withing the order of magnitude of our desired uncertainty in the calibration, the example clearly illustrates the need of including a bias correction step in our calibration methodology.

The ~~details about how this correction is made are fully explained in the support material. Summarizing, the procedure relies on using the standard deviation σ_ϵ between N~~ standard deviation σ_ϵ between N experimental retrievals of ~~C^0_Γ as an indicator of~~ $C^0_{\Gamma i}$ cannot be used directly as an estimation of uncertainty because the RCS distribution ~~'s shape . Then, we simulate an space of possible uncertainty sets,~~ shape is not Gaussian. The uncertainty introduced by this variability is studied by sampling a large set of possible RCS distributions based on our experimental configuration, ~~to generate the distribution Λ. Λ indicates how likely is a given bias value when the spread between N experiments is σ_ϵ (with a tolerance of 5%). The distribution Λ, shown in support material, is monomodal and assymetric. Because of this we use its median $\tilde{\Lambda}$ as the best bias correction estimator . In turn, the RMSE between Λ and~~ and selecting the candidates matching our observed spread σ_ϵ. This set provides an estimation of the expected bias correction $\tilde{\Lambda}$ ~~is used to estimate the~~ and of the effective RCS uncertainty σ_Λ. The uncertainty of the C^0_Γ estimator of Eq. (5) will correspond to the uncertainty of each $C^0_{\Gamma i}$ estimation propagated through the calculation of their average (terms $\sum\sigma_i^2/N^2$ and σ_T^2/N of Eq. (6a)) plus the effective RCS uncertainty σ_Λ. The details on how this estimator works and how the RCS distribution sampling is done are fully explained in Sect. S3 of the ~~bias correction, term which is later added to the uncertainty budget~~supplementary material.

6 Results

In 2018 we used the 20 m mast only, performing six iterations. For 2019 we did 10 iterations using the 10 m mast and 2 iterations with the 20 m mast. The distributions of C^0_Γ obtained in each iteration and experiment is shown in Fig. 10.

The radar hardware changed between 2018 and 2019 campaigns due to experiments required to retrieve the power transfer curve and perform maintenance operations. This implies that we cannot compare absolute calibration values between both campaigns. What remains valid is to compare properties such as the variability, and the results from both experiments of 2019.

In the results we can notice a difference in $C^0_{\Gamma i}$ spread when comparing the 10 and 20 m masts. The 6 iterations of 2018 (Fig. 10 (A)) have an spread of $\sigma_\epsilon = 0.33\ dB$, while the spread of the 10 iterations of 2019 is $0.11\ dB$ (Fig. 10 (B)). This happens because the 10 m mast has a motor on top which enables a much finer adjustment of the target position, improving the repeateability of the experiments.

There is also a small difference in the spread of the curves. The $C^0_{\Gamma i}$ values retrieved in experiment (B) have a smaller spread σ_i. This is because we took all the samples during one single night, with very clear conditions and an average wind speed below $1\ m/s$. A great advantage was the presence of the motor that enables target alignment in ≈ 5 minutes. Meanwhile, for experiment (A) curves were sampled during different days, because the 20 m mast setup requires more time to align (≈ 2 hours). The different conditions in each day led to a more varied shape in the retrieved curves. This effect is specially noticeable in experiment (C), where the iterations were performed during daytime, when atmospheric conditions are more dynamic, specially wind speed variability. The introduced variability was not fully compensated by our corrections and thus

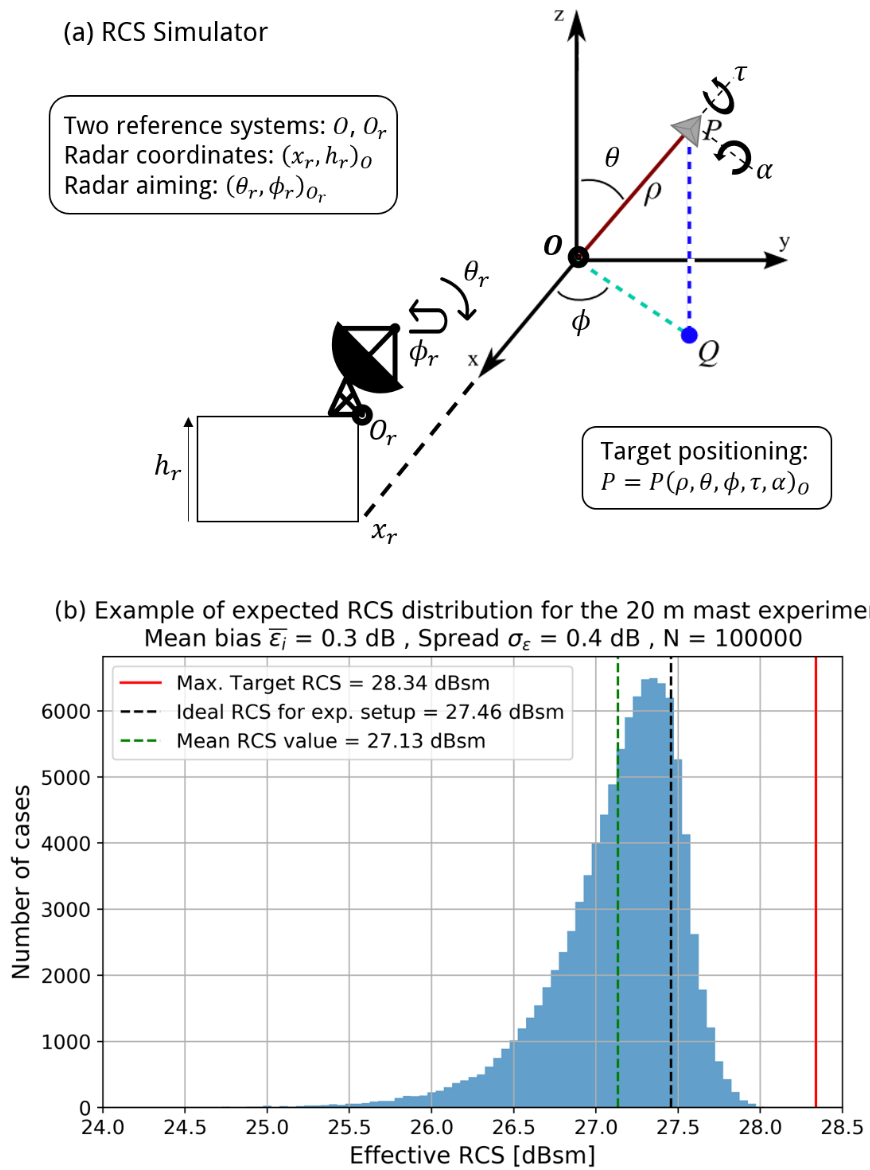

Figure 9. (a) Diagram of the RCS simulator illustrating its degrees of freedom. (b) Example of an effective RCS distribution obtained after 100 000 simulations with the uncertainty set specified in the text. The simulations are based on our 20 m mast setup. Bias is calculated substracting the ideal RCS by the mean RCS value. The example illustrates how the effective RCS will be, statistically, lower than the result expected from an ideally aligned setup.

bimodal distributions remained. However, individual spread is still small, within ≈ 0.1 dB, so we decided to accept these samples for calibration purposes.

To study the dependency of the bias correction on the amount of iterations we calculate the bias correction term $\tilde{\Lambda}$ and its uncertainty σ_Λ of experiments (A) and (B) with different amounts of repetitions. The order of the iterations used in each row match the sequential order indicated in Fig. 10. The results are shown in Table 1. For both cases we have the best estimate when we use all the samples available for each experiment, and thus we use this bias correction and uncertainty when computing the calibration coefficient.

For experiment (C) we followed a different approach. Because we only have two samples, the calculated $\sigma_\epsilon = 0.2\ dB$ is very likely to be underestimated. Consequently, and because the experimental procedure was identical to what was done in 2018, we assume our parameters σ_ϵ, $\tilde{\Lambda}$ and σ_Λ to be equal to the best estimation of experiment (A). This is possible because in our methodology we assume that the bias probability distribution of a given system is unique, even if it is unknown, and what is done by performing many iterations is to successively restrict the possible sets of uncertainties that can generate results consistent with the observations. This latter hypothesis is consistent with the decrease in ~~bias estimation uncertainty~~ uncertainty for the bias correction when increasing the amount of iterations.

Table 2 contains a summary of all known bias corrections and uncertainty contributions~~involved in the calculation of the calibration, as~~ , introduced in Sect. 4. With the afformentioned results, we use Eqs. (5)~~and (6) to calculate to calculate~~ , (2), (6a) and (6b) to estimate the RCS and Reflectivity calibration terms ~~$C_\Gamma(T)$ and $C_Z(T)$~~ $C_\Gamma(T,r)$ and $C_Z(T,r)$, alongside their uncertainty. ~~$C_Z(T)$~~ $C_Z(T,r)$ is calculated for the range resolution $\delta r = 12.5\ m$, which is the same mode used for target sampling. T is the radar amplifier temperature in ^oC ~~. Results are shown below~~and $f_{IF}(r)$ is the IF loss correction function.

- (A) 20 m mast - 2018:
  - ◇ ~~$C_\Gamma(T) = -275.6 + 0.093(T-26.5) \pm 0.3\ dB$~~ $C_\Gamma(T,r) = -275.6 + 0.093(T-26.5) + f_{IF}(r)\ [dB(m^{-2}AU^{-1})] \pm 0.3\ [dB]$
  - ◇ ~~$C_Z(T) = -191.5 + 0.093(T-26.5) \pm 0.3\ dB$~~ $C_Z(T,r) = -191.5 + 0.093(T-26.5) + f_{IF}(r)\ [dB(mm^6m^{-5}AU^{-1})] \pm 0.3\ [dB]$
  - ◇ $f_{IF}(r) = -7.62\cdot10^{-23}r^6 + 3.82\cdot10^{-19}r^5 - 4.69\cdot10^{-15}r^4 + 9.76\cdot10^{-11}r^3 - 4.99\cdot10^{-7}r^2 + 6.23\cdot10^{-4}r - 1.74\cdot10^{-1}\ [dB]$

- (B) 10 m mast - 2019:
  - ◇ ~~$C_\Gamma(T) = -274.4 + 0.093(T-26.5) \pm 0.9\ dB$~~ $C_\Gamma(T,r) = -274.4 + 0.093(T-26.5) + f_{IF}(r)\ [dB(m^{-2}AU^{-1})] \pm 0.9\ [dB]$
  - ◇ ~~$C_Z(T) = -190.3 + 0.093(T-26.5) \pm 0.9\ dB$~~ $C_Z(T,r) = -190.3 + 0.093(T-26.5) + f_{IF}(r)\ [dB(mm^6m^{-5}AU^{-1})] \pm 0.9\ [dB]$
  - ◇ $f_{IF}(r) = 4.87\cdot10^{-22}r^6 - 9.79\cdot10^{-18}r^5 + 6.35\cdot10^{-14}r^4 - 1.18\cdot10^{-10}r^3 - 1.16\cdot10^{-7}r^2 + 1.09\cdot10^{-4}r - 2.08\cdot10^{-2}\ [dB]$

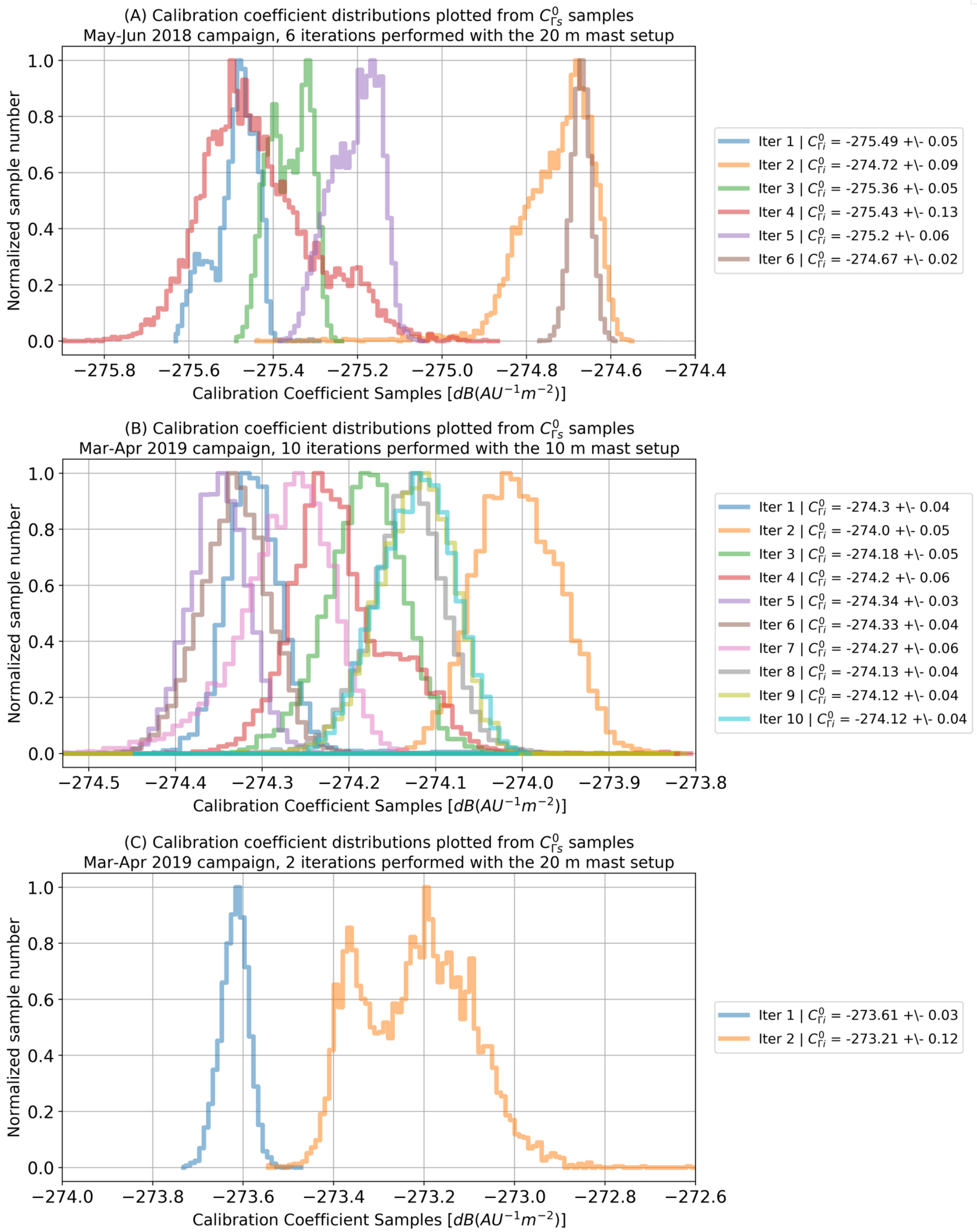

Figure 10. Calibration coefficient distributions obtained for (A) 2018 campaign using the 20 cm target on the 20 m mast, (B) 2019 campaign using the 10 cm target on the 10 m mast and (C) 2019 campaign with the 20 cm target on the 20 m mast.

Table 1. Bias correction $\tilde{\Lambda}$ and its uncertainty σ_Λ calculated using a different amount of iterations, for the experiments of 2018 and 2019 calibration campaigns (for ex. 3 iterations means we used iterations 1, 2 and 3 of the experiment). We include the average and spread σ_ϵ between the retrieved $C^0_{\Gamma i}$ for each case. ~~The~~ This variability σ_ϵ is introduced in the bias estimation ~~of~~ procedure to determine the bias correction $\tilde{\Lambda}$ and its uncertainty σ_Λ ~~depends on the amount of iterations N and their associated σ_ϵ value~~.

(A) 20 m mast 2018	Exp. Results		Bias Correction	
N^o of iterations	$\frac{1}{N}\sum C^0_{\Gamma i}$	σ_ϵ [dB]	$\tilde{\Lambda}$ [dB]	σ_Λ [dB]
2	-275.11	0.38	0.98	1.78
3	-275.19	0.33	0.65	0.86
4	-275.25	0.31	0.51	0.50
5	-275.24	0.28	0.40	0.33
6	-275.14	0.33	0.44	0.28
(B) 10 m mast 2019				
N^o of iterations				
2	-274.15	0.15	0.78	1.65
3	~~-274.15~~ -274.16	0.12	0.42	0.70
4	~~-274.16~~ -274.17	0.11	0.27	0.34
5	-274.20	0.12	0.24	0.20
6	~~-274.22~~ -274.23	0.12	0.22	0.13
7	-274.23	0.11	0.19	0.10
8	~~-274.21~~ -274.22	0.11	0.18	0.07
9	~~-274.20~~ -274.21	0.11	0.17	0.06
10	~~-274.19~~ -274.20	0.11	0.16	0.05
(C) 20 m mast 2019				
N^o of iterations				
2	-273.41	-	0.44	0.28

– (C) 20 m mast - 2019:

◇ ~~$C_\Gamma(T) = -273.9 + 0.093(T-26.5) \pm 0.4\, dB$~~ $C_\Gamma(T,r) = -273.9 + 0.093(T-26.5) + f_{IF}(r)\ [dB(m^{-2}AU^{-1})] \pm 0.4\ [dB]$

◇ ~~$C_Z(T) = -189.8 + 0.093(T-26.5) \pm 0.4\, dB$~~ $C_Z(T,r) = -189.8 + 0.093(T-26.5) + f_{IF}(r)\ [dB(mm^6 m^{-5} AU^{-1})] \pm 0.4\ [dB]$

◇ $f_{IF}(r) = 4.87\cdot 10^{-22} r^6 - 9.79\cdot 10^{-18} r^5 + 6.35\cdot 10^{-14} r^4 - 1.18\cdot 10^{-10} r^3 - 1.16\cdot 10^{-7} r^2 + 1.09\cdot 10^{-4} r - 2.06\cdot 10^{-2}\ [dB]$

~~Finally, we performed~~ These results enable the analysis of relative uncertainty contributions from different sources, however the total calibration uncertainty may be underestimated. As indicated in Sects. 4 and 5, some bias terms remain unknown.

Table 2. Summary of all known corrections and uncertainty contributions in the calculation of $C_\Gamma(T)$ $C_\Gamma(T,r)$. The absolute correction terms have a sign associated to with the direction in which they impact the final calibration calculation. For the receiver compression correction we present the average magnitude and for the temperature correction we present the range of possible values.

Absolute Corrections	Term [dB]	(A) 20 m mast 2018	(B) 10 m mast 2019	(C) 20 m mast 2019
Compression	Fig. 4 (a)	-0.3 in avg.	-0.2 in avg.	-0.3 in avg.
Partial Antenna Overlap	L_o $L_o(r_0)$	-0.02	-0.08	-0.02
Temp. Corr. ($T_0 = 26.5$ °C)	$n(T - T_0)$	within ±0.6	within ±0.6	within ±0.6
Misalignment Bias	$\tilde{\Lambda}$	-0.44	-0.16	-0.44
IF loss correction	$f_{IF}(r)$	$\leq$ |0.6|	$\leq$ |0.9|	$\leq$ |0.9|
Uncertainty Sources	**Term [dB]**			
$C^0_{\Gamma i}$ estimation	$\sqrt{\frac{1}{N^2}\sum\sigma_i^2}$	0.03	0.01	0.09
Temp. Corr. in $C^0_{\Gamma i}$ retrievals	$\frac{\sigma_T}{\sqrt{N}}$	0.05	0.04	0.09
Temp. Corr. in $C_\Gamma(T)$, $C_Z(T)$ $C_\Gamma(T,r)$, $C_Z(T,r)$	σ_T	0.13	0.13	0.13
Signal to Clutter Ratio	σ_{SCR}	0.09	0.93	0.09
Bias Correction	σ_Λ	0.28	0.05	0.28
IF loss correction	σ_{IF}	≤ 0.04	≤ 0.04	≤ 0.04
Final Total Calibration Uncertainty	δC_Γ; δC_Z	0.33	0.94	0.35

Specifically, target physical RCS must be measured in an echo chamber to improve the misalignment bias estimation. In addition, the method to characterize antenna alignment must be improved to determine if there is a need for an additional distance correction function (Sect. 5.2). The uncertainty of these retrievals will impact the total uncertainty value, however, it is possible to quantify this effect through the terms σ_{Γ_0} and σ_A of Eq. (6b).

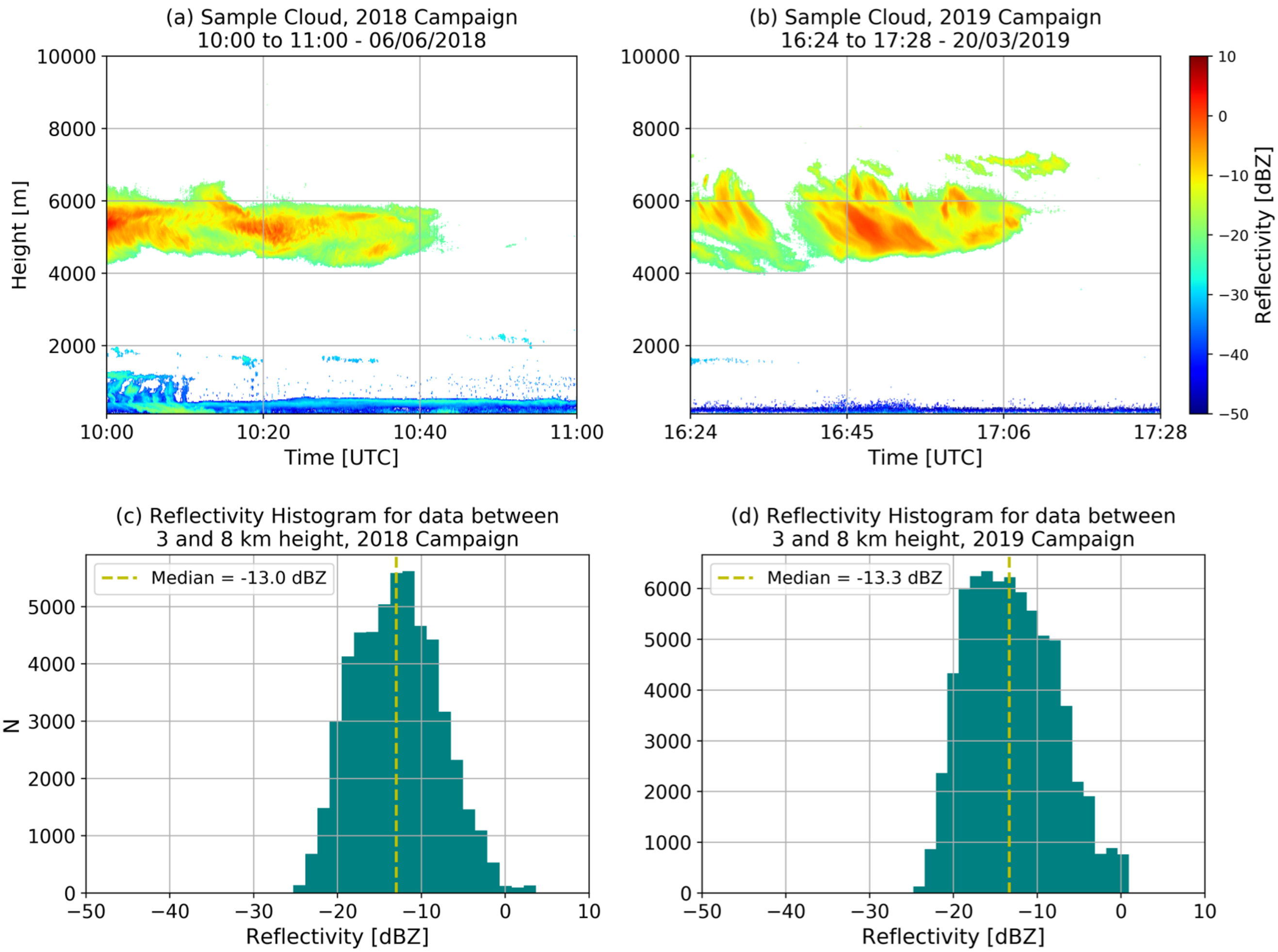

Figure 11. Altostratus cloud sampled during 2018 (a) and 2019 campaigns (b). Lower reflectivities are easier to capture at lower altitudes because of lower distance and attenuation losses (Eq. (4b)). In the altostratus reflectivity histograms (c) and (d) we observe that for both campaigns measurements are within the ranges reported in literature.

To finalize, we perform a test of the calibration results by measuring a altostratus cloud in both campaigns (Fig. 11). The sampling was done with the 25 m resolution, and thus 6 dB had to be substracted from the ~~$C_Z(T)$~~ $C_Z(T,r)$ calibration calculated for the 12.5 m resolution. In this correction, 3 dB come from the change in the distance resolution term δr (Eq. (4a)), and the other 3 dB are substracted to compensate the additional digital gain k_d coming from doubling the amount of points in the chirp fourier transform (Delanoë et al., 2016). A Signal to Noise Ratio threshold of 8 dB is used to remove noise samples. We observe that for both campaigns the reflectivity measured in altostratus cloud is within -30 - 0 dBZ, which are typical values reported in literature (Uttal and Kropfli, 2001).

7 Conclusions

This study presents a cloud radar calibration method that is based on cloud radar power signal backscattered from a reference reflector. We study the validity of the method and variability of the results by performing measurements in two experimental setups and analyzing the associated results. In the first experimental setup we use a scanning BASTA-Mini W-band cloud radar, that aims towards a 20-cm triangular trihedral target installed at the top of a 20-m mast, located 376.5 m from the radar. For the second experimental setup, we use the same radar, aimed towards a 10-cm triangular trihedral target mounted on a pan–tilt motor at the top of a 10-m mast. The mast is located 196 m from the radar.

The first consideration in the design of the experimental setup is the need to avoid excessive compression or saturation in the radar receiver. This must be checked before any calibration attempt by comparing measurements of radar backscattered power with the radar receiver power transfer curve. In both our setups we find losses due to compression on the order of $0.2 \sim 0.3$ dB. There is a compensating effect between target RCS and radar-to-target distance (Eq. 1b). Since the compression effect is small, we correct it using our receiver power transfer curve. However, in cases where the radar is operating close to saturation, or when compression effects are larger than the calibration uncertainty goal, it is advisable to compensate by reducing target size or by positioning the target ~~further~~ farther away from the radar.

Secondly, the reflector must be positioned far enough from the radar to be outside the antennas near-field distance and to ensure that the received power has low antenna-overlap losses. The BASTA-Mini cloud radar has a Fraunhofer near-field distance of 50 m. The estimated maximum overlap loss is less than 0.1 dB for the closest (10-m) mast setup. Thus we conclude that the target positioning is far enough for both setups.

Thirdly, the experimental setup should strive to reduce clutter in the radar measurements. This can be achieved by operating in an open field that is several hundred meters in length and free of trees or other signal-inducing obstacles. It is also advisable to perform radar measurements under clear conditions, without fog or rain, with wind speed below 1 ms^{-1} and low turbulence.

Next, the proposed calibration method requires performing several iterations in the same setup configuration. In each iteration the setup is first realigned, followed by approximately one hour of sampling of the reference reflector backscattered power. The sampled power is then corrected for compression effects, incomplete antenna overlap, variations in radar gain due to temperature and atmospheric attenuation, before being used to estimate a RCS calibration term value. Once all iterations are completed, the final RCS and Equivalent Reflectivity calibration terms can be computed with their respective uncertainties.

Iterations are necessary because they enable the quantification of bias introduced by inevitable system misalignment. Our experiments indicate that, for our setup, at least 5 iterations are necessary to reach convergence in the calculation of bias and uncertainty associated with misalignment. We find a bias correction of $\approx 0.4 \pm 0.3$ dB for the 20-m mast, and of $\approx 0.2 \pm 0.1$ dB for the 10-m mast. This difference can be explained by the more precise alignment attainable with the pan–tilt motor installed on the 10 m mast.

Calibration is also impacted by changes in the gain of radar components associated with internal temperature variations. For the radar used in our experiment, these changes reach up to ± 0.6 dB. Our experiments enabled us to retrieve a correction function for the temperature dependence and to reduce the temperature uncertainty contribution to $\sigma_T = 0.13\ dB$. This result indicates that lower calibration uncertainties can be achieved by studying temperature effects, especially for solid state radars.

Another necessary consideration is the inclusion of gain variations with distance, introduced by frequency dependent losses in the IF of the radar receiver. We found calibration variations with distance up to $0.9\ dB$ for the 2019 campaign. Therefore, characterizing the IF loss is a necessary step to validate the calibration results for all ranges.

Our analyses reveal that the predominant sources of uncertainty in our experimental setups are due to levels of clutter and alignment precision. These two effects have different magnitudes in our two experimental setups (10-m and 20-m masts). The 20-m mast setup uncertanty is limited by the uncertainty contribution of the alignment bias estimation $\sigma_\Lambda = 0.28\ dB$. The 10-m mast setup uncertainty is limited by the uncertainty contribution of the signal-to-clutter ratio $\sigma_{SCR} = 0.9\ dB$. This result reveals that there is a tradeoff between better target alignment and additional clutter introduced by the alignment motor.

The complete uncertainty budget enables us to conclude that the proposed calibration method can yield uncertainties as low as $0.3\ dB$ with our current equipment. This result was obtained using the 20-cm target on the 20 m mast during the 2018 experiment, where six target sampling iterations were performed. Additionally, in 2019 two completely different calibration setups were used with the same radar hardware, and in both cases, we obtain the same calibration result within uncertainty bounds.

Finally, because of cloud radar hardware evolutions in the fall of 2018, the calibration coefficients found in May 2018 and March 2019 differ by $1.2\ dB$. We compare cloud radar measurements of altrostratus clouds performed in May 2018 and March 2019. The reflectivity distributions of the two events are consistent and compatible with values previously registered in literature. The two distributions yield median values that differ by $0.3\ dB$.

For future work we envisage to develop a technological solution to allow target orientation without introducing additional clutter. Another interesting prospect is to improve the accuracy of the radar positioner, to enable direct retrieval of antenna pattern or target RCS directly with the radarinstead of relying on laboratory measurements, following the method proposed by Garthwaite et al. (2015). This retrieval would improve bias correction arising from parallax errors. We also plan to perform an echo chamber characterization of our reference targets, to remove any possible bias caused by manufacturing imprecision and to improve the estimation of our misalignment bias correction.

Further, there is ongoing research on calibration and antenna pattern characterization methods based on reference targets held by Unmanned Aerial Vehicles (UAVs) (Yin et al., 2019)(Duthoit et al., 2017; Yin et al., 2019). Since the underlying principle is the same, most considerations written here should be directly applicable in these new experiments. Here the UAV takes the role of the mast, holding the reflector (usually a sphere), and therefore it is important to characterize the UAV RCS and verify that it does not interfere with the experiment. The main difference would be in the procedure necessary to estimate bias, because the reference target (usually a sphere) will be always moving due to wind. Here an adaptation of the effective RCS simulator would be necessary to account for the target type and different alignment protocol.

Author contributions.

All authors contributed to the planning of the campaigns and the design of the calibration experiments.

Author Julien Delanoë was responsible of radar installation and operation.

Authors Jean-Charles Dupont and Felipe Toledo worked in the preparation, development and operation of the necessary infrastructure for the experiments.

Authors Julien Delanoë and Felipe Toledo retrieved the Power Transfer Curve of the Radar Receiver.

Data analysis and the establishment of the calibration methodology presented in the paper was done by Felipe Toledo.

Authors Martial Haeffelin and Felipe Toledo worked in defining the paper structure and content.

Authors Felipe Toledo and Susana Jorquera worked in developing the method to retrieve the IF correction function, and in its calculation.

All authors reviewed the paper.

Competing interests.

Author Felipe Toledo has received research funding from Company Meteomodem.

Acknowledgements. The authors would like to acknowledge Johan Parra, Patricia Delville, Cristophe Boitel and Marc-Antoine Drouin from the SIRTA Atmospheric Observatory for their assistance in the execution of the field experiments. This acknowledgement is extended to Razvan Pirloaga and Dragos Ene from the INOE Institute, Romania. We would also like to thank Christophe Le Gac, Fabrice Bertrand ~~,~~ and Jean-Paul Vinson ~~and Susana Jorquera~~ from the LATMOS Laboratory, France, for their scientific and technical collaboration.

We would like to acknowledge the two reviewers for their expert comments which enabled us to improve the proposed calibration method.

We also acknowledge the french *Association Nationale de la Recherche* (ANRT) and the company Meteomodem for their contribution in the funding of this work. Finally, we state this work is part of the ACTRIS-2 project and has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654109.

Symbols

Symbol	Description	Units
$C_\Gamma(T,r)$	RCS Calibration Term	$dB(AU^{-1}m^{-2})$
$C_\Gamma(T,r_0)$	RCS Calibration Term at the target position r_0	$dB(AU^{-1}m^{-2})$
C_Γ^0	RCS Calibration Coefficient	$dB(AU^{-1}m^{-2})$
$C_{\Gamma s}^0$	Single sample of the Calibration Coefficient C_Γ^0	$dB(AU^{-1}m^{-2})$
$C_Z(T,r)$	Radar Equivalent Reflectivity Calibration term	$dB(mm^6m^{-5}AU^{-1})$
δC_Γ	RCS calibration uncertainty	dB

Symbol	Description	Units
δC_Z	Reflectivity calibration uncertainty	dB
$f_{IF}(r)$	IF loss correction function	dB
$\Gamma(r)$	Radar Cross Section of reflections at a distance r	$dBsm$
Γ_0	Radar Cross Section of the reference target	$dBsm$
k_d	Digital gain of the radar receiver	
$\tilde{\Lambda}$	Misalignment bias correction	dB
λ	Radar carrier wavelength	m
N	Number of iterations performed in a calibration experiment	
$P_r(r_0)$	Power received from the target position r_0	$dB(AU)$
$P_r(r)$	Power received from distance r	$dB(AU)$
p_t	Radar transmitted power	W
r	Distance from the radar	m
r_0	Distance between the reference target and the radar	m
σ_A	Uncertainty of the antenna properties (beam shape and alignment)	dB
σ_ϵ	Standard deviation between the N mean RCS Calibration Coefficients $C^0_{\Gamma i}$, used to calculate the bias correction	dB
σ_{Γ_0}	Uncertainty of the reference target RCS	dB
σ_i	Standard deviation of the RCS Calibration Coefficient samples $C^0_{\Gamma s}$ for iteration i	dB
σ_{IF}	Uncertainty of the IF loss correction function	dB
σ_Λ	Uncertainty of the misalignment bias correction	dB
σ_{SCR}	Uncertainty introduced by clutter at the target position	dB
σ_T	Uncertainty of the temperature correction function	dB
θ	Antenna beamwidth	rad
Z_e	Radar Equivalent Reflectivity	dBZ

References

Anagnostou, E. N., Morales, C. A., and Dinku, T.: The Use of TRMM Precipitation Radar Observations in Determining Ground Radar Calibration Biases, Journal of Atmospheric and Oceanic Technology, 18, 616–628, https://doi.org/10.1175/1520-0426(2001)018<0616:TUOTPR>2.0.CO;2, https://doi.org/10.1175/1520-0426(2001)018<0616:TUOTPR>2.0.CO;2, 2001.

Atlas, D.: RADAR CALIBRATION, Bulletin of the American Meteorological Society, 83, 1313–1316, https://doi.org/10.1175/1520-0477-83.9.1313, https://doi.org/10.1175/1520-0477-83.9.1313, 2002.

Bergada, M., Sekelsky, S. M., and Li, L.: External calibration of millimeter-wave atmospheric radar systems using corner reflectors and spheres, cal, 3, 6, 2001.

Boers, R., Baltink, H. K., Hemink, H. J., Bosveld, F. C., and Moerman, M.: Ground-Based Observations and Modeling of the Visibility and Radar Reflectivity in a Radiation Fog Layer, Journal of Atmospheric and Oceanic Technology, 30, 288–300, https://doi.org/10.1175/JTECH-D-12-00081.1, https://doi.org/10.1175/JTECH-D-12-00081.1, 2013.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., et al.: Clouds and aerosols, in: Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, pp. 571–657, Cambridge University Press, 2013.

Bringi, V. N. and Chandrasekar, V.: Polarimetric Doppler weather radar: principles and applications, Cambridge university press, 2001.

Brooker, G.: Sensors and signals, Citeseer, Australian Centre for Field Robotics, Rose St Building (J04), University of Sydney, Australia, 2006.

Chandrasekar, V., Baldini, L., Bharadwaj, N., and Smith, P. L.: Calibration procedures for global precipitation-measurement ground-validation radars, URSI Radio Science Bulletin, 2015, 45–73, https://doi.org/10.23919/URSIRSB.2015.7909473, 2015.

Delanoë, J., Protat, A., Vinson, J.-P., Brett, W., Caudoux, C., Bertrand, F., Parent du Chatelet, J., Hallali, R., Barthes, L., Haeffelin, M., et al.: Basta: a 95-GHz fmcw doppler radar for cloud and fog studies, Journal of Atmospheric and Oceanic Technology, 33, 1023–1038, 2016.

Doerry, A. W. and Brock, B. C.: Radar cross section of triangular trihedral reflector with extended bottom plate, Sandia Report, Sandia National Laboratory, 2009.

Doviak, R. J. and Zrnić, D. S.: Doppler Radar and Weather Observations, Mineola, New York: Dover Publications, INC, 2006.

Dupont, J.-C., Haeffelin, M., Protat, A., Bouniol, D., Boyouk, N., and Morille, Y.: Stratus–fog formation and dissipation: a 6-day case study, Boundary-layer meteorology, 143, 207–225, 2012.

Dupont, J.-C., Haeffelin, M., Wærsted, E., Delanoë, J., Renard, J.-B., Preissler, J., and O'dowd, C.: Evaluation of Fog and Low Stratus Cloud Microphysical Properties Derived from In Situ Sensor, Cloud Radar and SYRSOC Algorithm, Atmosphere, 9, 169, 2018.

Duthoit, S., Salazar, J. L., Doyle, W., Segales, A., Wolf, B., Fulton, C., and Chilson, P.: A new approach for in-situ antenna characterization, radome inspection and radar calibration, using an Unmanned Aircraft System (UAS), in: 2017 IEEE Radar Conference (RadarConf), pp. 0669–0674, 2017.

Ewald, F., Groß, S., Hagen, M., Hirsch, L., Delanoë, J., and Bauer-Pfundstein, M.: Calibration of a 35 GHz airborne cloud radar: lessons learned and intercomparisons with 94 GHz cloud radars, Atmospheric Measurement Techniques, 12, 1815–1839, https://doi.org/10.5194/amt-12-1815-2019, https://www.atmos-meas-tech.net/12/1815/2019/, 2019.

Fox, N. I. and Illingworth, A. J.: The retrieval of stratocumulus cloud properties by ground-based cloud radar, Journal of Applied Meteorology, 36, 485–492, 1997.

Garthwaite, M. C., Nancarrow, S., Hislop, A., Thankappan, M., Dawson, J. H., and Lawrie, S.: Design of radar corner reflectors for the Australian Geophysical Observing System, Geoscience Australia, 3, 2015.

Gaussiat, N., Sauvageot, H., and Illingworth, A. J.: Cloud liquid water and ice content retrieval by multiwavelength radar, Journal of Atmospheric and Oceanic Technology, 20, 1264–1275, 2003.

Haeffelin, M., Barthès, L., Bock, O., Boitel, C., Bony, S., Bouniol, D., Chepfer, H., Chiriaco, M., Cuesta, J., Delanoë, J., et al.: SIRTA, a ground-based atmospheric observatory for cloud and aerosol research, in: Annales Geophysicae, vol. 23, pp. 253–275, 2005.

Haynes, J. M., L'Ecuyer, T. S., Stephens, G. L., Miller, S. D., Mitrescu, C., Wood, N. B., and Tanelli, S.: Rainfall retrieval over the ocean with spaceborne W-band radar, Journal of Geophysical Research: Atmospheres, 114, 2009.

Hogan, R. J., Jakob, C., and Illingworth, A. J.: Comparison of ECMWF winter-season cloud fraction with radar-derived values, Journal of Applied Meteorology, 40, 513–525, 2001.

Liebe, H. J.: MPM—An atmospheric millimeter-wave propagation model, International Journal of Infrared and millimeter waves, 10, 631–650, 1989.

Liebe, H. J., Manabe, T., and Hufford, G. A.: Millimeter-wave attenuation and delay rates due to fog/cloud conditions, IEEE Transactions on Antennas and Propagation, 37, 1617–1612, 1989.

Meissner, T. and Wentz, F. J.: The complex dielectric constant of pure and sea water from microwave satellite observations, IEEE Transactions on Geoscience and Remote Sensing, 42, 1836–1849, https://doi.org/10.1109/TGRS.2004.831888, 2004.

Mülmenstädt, J. and Feingold, G.: The radiative forcing of aerosol–cloud interactions in liquid clouds: wrestling and embracing uncertainty, Current Climate Change Reports, 4, 23–40, 2018.

Myhre, G., Shindell, D., Bréon, F. M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., and Lee, D.: Mendoza, 25 B, Nakajima, T., Robock, A., Stephens, G., Takemura, T., and Zhang, H. Anthropogenic and natural radiative forcing. In: Stocker, T., Qin, D., Plattner, GK, Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P.(eds) Climate Change, 2013.

Pappalardo, G.: ACTRIS Aerosol, Clouds and Trace Gases Research Infrastructure, EPJ Web Conf., 176, 09 004, https://doi.org/10.1051/epjconf/201817609004, https://doi.org/10.1051/epjconf/201817609004, 2018.

Pozar, D. M.: Microwave engineering, John Wiley & Sons, 2009.

Protat, A., Bouniol, D., Delanoë, J., O'Connor, E., May, P. T., Plana-Fattori, A., Hasson, A., Görsdorf, U., and Heymsfield, A. J.: Assessment of CloudSat reflectivity measurements and ice cloud properties using ground-based and airborne cloud radar observations, Journal of Atmospheric and Oceanic Technology, 26, 1717–1741, 2009.

Protat, A., Bouniol, D., O'Connor, E. J., Klein Baltink, H., Verlinde, J., and Widener, K.: CloudSat as a Global Radar Calibrator, Journal of Atmospheric and Oceanic Technology, 28, 445–452, https://doi.org/10.1175/2010JTECHA1443.1, https://doi.org/10.1175/2010JTECHA1443.1, 2011.

Richardson, M.: Fundamentals of the discrete fourier transform, Sound and vibration magazine, pp. 40–46, 1978.

Sassen, K.: Ice cloud content from radar reflectivity, Journal of climate and applied meteorology, 26, 1050–1053, 1987.

Scolnik, M. I.: Radar Handbook, Mc Craw-Hill, 3 edn., 2000.

Sekelsky, S. M. and Clothiaux, E. E.: Parallax Errors and Corrections for Dual-Antenna Millimeter-Wave Cloud Radars, Journal of Atmospheric and Oceanic Technology, 19, 478–485, https://doi.org/10.1175/1520-0426(2002)019<0478:PEACFD>2.0.CO;2, https://doi.org/10.1175/1520-0426(2002)019<0478:PEACFD>2.0.CO;2, 2002.

Uttal, T. and Kropfli, R. A.: The Effect of Radar Pulse Length on Cloud Reflectivity Statistics, Journal of Atmospheric and Oceanic Technology, 18, 947–961, https://doi.org/10.1175/1520-0426(2001)018<0947:TEORPL>2.0.CO;2, https://doi.org/10.1175/1520-0426(2001)018<0947:TEORPL>2.0.CO;2, 2001.

Wærsted, E. G., Haeffelin, M., Dupont, J.-C., Delanoë, J., and Dubuisson, P.: Radiation in fog: quantification of the impact on fog liquid water based on ground-based remote sensing, Atmospheric Chemistry and Physics, 17, 10 811–10 835, https://doi.org/10.5194/acp-17-10811-2017, https://www.atmos-chem-phys.net/17/10811/2017/, 2017.

Wærsted, E. G., Haeffelin, M., Steeneveld, G.-J., and Dupont, J.-C.: Understanding the dissipation of continental fog by analysing the LWP budget using idealized LES and in situ observations, Quarterly Journal of the Royal Meteorological Society, 145, 784–804, https://doi.org/10.1002/qj.3465, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3465, 2019.

Yau, M. K. and Rogers, R. R.: A short course in cloud physics, Elsevier, 1996.

Yin, J., Hoogeboom, P., Unal, C., Russchenberg, H., Van der Zwan, F., and Oudejans, E.: UAV-Aided Weather Radar Calibration, IEEE Transactions on Geoscience and Remote Sensing, 57, 10 362–10 375, https://doi.org/10.1109/TGRS.2019.2933912, 2019.