

Article

Hierarchical Bayesian Data Analysis in Radiometric SAR System Calibration: A Case Study on Transponder Calibration with RADARSAT-2 Data

Björn J. Döring *, Kersten Schmidt, Matthias Jirousek, Daniel Rudolf, Jens Reimann, Sebastian Raab, John Walter Antony and Marco Schwerdt

Microwaves and Radar Institute, German Aerospace Center (DLR), Oberpfaffenhofen, D-82234 Weßling, Germany; E-Mails: kersten.schmidt@dlr.de (K.S.); matthias.jirousek@dlr.de (M.J.); daniel.rudolf@dlr.de (D.R.); jens.reimann@dlr.de (J.R.); sebastian.raab@dlr.de (S.R.); john.walterantony@dlr.de (J.W.A.); marco.schwerdt@dlr.de (M.S.)

* Author to whom correspondence should be addressed; E-Mail: bjoern.doering@dlr.de; Tel.: +49-8153-28-2246; Fax: +49-8153-28-1449.

Received: 19 October 2013; in revised form: 20 November 2013 / Accepted: 28 November / Published: 4 December 2013

Abstract: A synthetic aperture radar (SAR) system requires external absolute calibration so that radiometric measurements can be exploited in numerous scientific and commercial applications. Besides estimating a calibration factor, metrological standards also demand the derivation of a respective calibration uncertainty. This uncertainty is currently not systematically determined. Here for the first time it is proposed to use hierarchical modeling and Bayesian statistics as a consistent method for handling and analyzing the hierarchical data typically acquired during external calibration campaigns. Through the use of Markov chain Monte Carlo simulations, a joint posterior probability can be conveniently derived from measurement data despite the necessary grouping of data samples. The applicability of the method is demonstrated through a case study: The radar reflectivity of DLR's new C-band Kalibri transponder is derived through a series of RADARSAT-2 acquisitions and a comparison with reference point targets (corner reflectors). The systematic derivation of calibration uncertainties is seen as an important step toward traceable radiometric calibration of synthetic aperture radars.

Keywords: synthetic aperture radar; radiometric calibration; external calibration; Bayesian data analysis; transponder

1. Introduction

A synthetic aperture radar (SAR) system is a measurement system that acquires measurement data for earth-observation applications. Many applications like soil moisture [1] or forest biomass estimation [2] require absolutely radiometrically calibrated data products. The absolute radiometric calibration ensures that systematic effects have been characterized and are compensated so that the derived radar ground reflectivity appears to be independent of the measurement system. Not until then can the data be used for physical parameter inversion modeling.

The necessary radiometric calibration is a two-step process and is divided into relative and absolute calibration [3,4]. During relative calibration, the system transfer function, antenna pattern, instrument drift (due to temperature and time), antenna mis-pointing and mis-polarization, SAR mode (Stripmap, ScanSAR, *etc.*), atmosphere, and processor effects are characterized and compensated. After relative calibration, the mapped reflectivities are still only given as system indications, *i.e.*, as digital numbers derived during processing. The digital numbers are then related to the definition of a physical unit through absolute calibration. Absolute calibration is, for practical reasons and an expected higher calibration quality, executed as an *external* calibration for space-borne SAR systems [3]. Reference point targets (typically corner reflectors or active radar calibrators [3–6]) with known radar reflectivities are distributed within an imaged scene, and their backscatters are related to the indications from the SAR system through a proportionality constant, often called calibration factor K . Usually the external calibration is repeated for a number of acquisitions, each using a number of reference point targets, in order to reduce random effects through averaging.

1.1. Problem Statement

The result of a measurement is only meaningful if, besides the estimate of the value of the measurand, a statement about the *uncertainty* of the estimate is given. Uncertainty is defined here as “the parameter which characterizes the dispersion of the values that could reasonably be attributed to the measurand” [7]. Stating uncertainties is the prerequisite for measurement results which are traceable to national standards. Metrological traceability is necessary for the comparability of measurement results, the foundation for quantitative scientific and commercial applications.

The final, combined measurement uncertainty results from several uncertainty contributions, and the calibration uncertainty is one of them. If a SAR system is absolutely calibrated, the uncertainty with which the radar reflectivity of the reference point target is known has a direct impact on the combined uncertainty. Another source of uncertainty is the estimation of the calibration factor from external calibration measurements. Specifically, this second source of uncertainty is addressed in this paper.

Currently, the derivation of measurement uncertainties in radiometric SAR system calibration is not systematically conducted. To the best of our knowledge, no operator of a previous or current space-borne SAR system claims to provide traceable radiometric calibration, and the quoted measurement uncertainties in the respective data product specifications or other publications [5,6,8–11] remain questionable. At the moment, a comprehensive uncertainty analysis, preferably in accordance with the ISO *Guide to the expression of uncertainty in measurement* (GUM) [7], is lacking. This diminishes the scientific integrity of the data products and hinders further adoption of radiometric

measurement data for advanced quantitative analysis and physical parameter inversion modeling, within and across SAR missions.

1.2. Objective, Approach, and Paper Structure

One piece of the puzzle of achieving traceable radiometric SAR system calibration is the derivation of the calibration factor from external calibration acquisitions. We propose for the first time to apply Bayesian statistical data analysis and hierarchical modeling to estimate parameters like the absolute calibration factor *including uncertainties and confidence intervals* from SAR images. The proposed method is also applicable for similar data analysis tasks in radiometric calibration.

Bayesian data analysis has been proven advantageous in many parameter estimation problems [12,13], but to the best of our knowledge so far it has not yet been applied for absolute radiometric SAR calibration. Bayesian analysis lends itself well to hierarchical data modeling [14], and is therefore a very good fit for the estimation problems that one faces during an external SAR calibration campaign. A general but short review of the methodology and an outline of its application in the domain of absolute radiometric calibration is given in Section 2.

The suitability of the proposed method is shown through a case study in Sections 3 (campaign setup) and 4 (data analysis). The objective of the case study is conceptually identical to the derivation of the absolute calibration factor, although here the reflectivity of a point target shall be accurately determined. The presented case study is based on an external calibration campaign which was executed in April 2013. 15 corner reflectors were deployed as reference targets, and the reflectivity of DLR's new C-band Kalibri transponder prototype was derived from a series of eight data acquisitions from the Canadian RADARSAT-2 SAR system.

Finally, the proposed approach is further discussed in Section 5 and conclusions are given in Section 6.

1.3. Note on Point Target RCS versus ERCS

This paper adopts the reasoning laid out in [15]. According to it, a reference point target's brightness in a SAR image is not generally proportional to its radar cross section (RCS), which is a body property depending on frequency and incidence angle. Rather, a target's brightness in a SAR image depends on the measured *complex* amplitudes and the *weighted averaging* (over the chirp pulse bandwidth and the azimuth angles) executed by the SAR processor. This other quantity is called *equivalent radar cross section*, or ERCS.

Targets whose ERCS is accurately known can be taken as a reference to calibrate SAR systems. The aim of the presented case study is to derive the ERCS of DLR's next-generation C-band Kalibri transponder prototype so that in principle it can be taken as a calibration normal for subsequent SAR calibration campaigns.

In order to avoid confusion between the symbols for (E)RCS and standard deviation, the (E)RCS will not be denoted with the customary letter σ but with its alternative form ζ .

2. Methodology for Parameter Estimation from SAR Data

Bayesian statistics is, like classical (frequentist) statistics, a well established field with applications in many scientific areas. It is extensively covered in the literature (see for example [12,13,16]), and the following aims to only sketch basic principles and to highlight its usefulness in the frame of absolute radiometric SAR system calibration.

2.1. Introduction to Bayesian Statistics and Numerical Methods

In Bayesian statistics, all unknown quantities are handled as random variables and are described by probability distributions, which in turn describe a measure of the state of knowledge of the parameter's value. In Bayesian analysis, the probability function of a parameter (prior) is updated by incorporating new information, *i.e.*, data, to derive a posterior probability function.

If a population parameter (e.g., the case study transponder ERCS) is called θ , then it can be described by a probability distribution function $p(\theta)$. Even before a measurement is done, a certain, subjective, (prior) probability density can be attributed to the parameter. In case of the transponder's ERCS, one might claim that the parameter needs to be positive and that it certainly will not be above 100 dBm². If now some new data y (e.g., from the case study RADARSAT-2 campaign) becomes available, one wants to update one's believe on θ given data y . This can be written as $p(\theta|y)$ where $p(\cdot|\cdot)$ describes a conditional probability. The notation follows [12]. Now, Bayes' rule allows to compute the posterior probability $p(\theta|y)$ from a prior probability $p(\theta)$ and a likelihood function $p(y|\theta)$ [12]:

$$p(\theta|y) = \frac{p(\theta)p(y|\theta)}{p(y)} \quad (1)$$

The likelihood function $p(y|\theta)$ describes how likely the data is, given the parameter θ .

Describing parameters by probability distributions is a natural fit when not only the best estimate of a parameter needs to be stated, but also a confidence interval needs to be quantified, as is the case for calibration. Deriving a confidence interval from a distribution is straight forward for any kind of distribution, *i.e.*, the analysis is not limited to Gaussian distributions.

Deriving posterior distributions is in practice mostly achieved through numerical methods, which allow to consider more complex problems and arbitrary distributions. The simulation method used in the case study is the Markov chain Monte Carlo (MCMC) approach [12]. The posterior distribution is approximated by sequentially drawing samples from approximate iterative intermediate distributions until the simulation converges to the target distribution. Several software packages exist which simplify computations; here the PyMC library is used [17].

2.2. Hierarchical Models

An external radiometric calibration campaign results in a pool of data samples (see [5] for an example), one per reference point target ERCS per SAR image. Depending on the research question, the data often needs to be grouped to estimate group-level parameters (for compensation) or to select group-level models (for validation). Typical questions are:

- What is the best estimate of the calibration factor (and its respective confidence interval) if several types of reference point targets (*i.e.*, transponders and corners of different sizes) with different ERCS' and stabilities are deployed?

Solving this problem with classical (frequentist) statistics would require to estimate the population mean of each group, and deriving the calibration factor after ERCS compensation between groups. The information on the variance within each group is lost, and a reliable statement of the final uncertainty or confidence interval on the estimated calibration factor is difficult to achieve. With hierarchical Bayesian modeling though, the variance within each group (target type) and the variance across all target types can be derived simultaneously because group and total dispersion are handled within a joint probability model.

- Is there a significant systematic dependence on the chosen antenna beam (or near/far range, left/right looking geometries, or ascending/descending orbits) for radiometric measurements?

Once again the same set of data samples as before should be grouped, but this time by antenna beam (or near/far range, left/right looking acquisitions, or ascending/descending orbits). For each group, a posterior distribution for the respective calibration factor can now be derived. Comparing the different posterior distributions allows to conclude if a significant radiometric inter-beam offset exists.

- For a check on plausibility: Is the ERCS of one of the reference point targets systematically different from the others? (Here repeated overpasses over the same set of targets is assumed.)

In order to answer this question, the overpass-dependent effect of the SAR system and the atmosphere should be modeled out of the analysis. This can be done by grouping the samples according to overpass and target ID. All target samples of one overpass can be used to compensate for SAR system and atmospheric effects, and in a second step the group ERCS of each target can be determined.

In fact, all questions previously raised can be conveniently answered by setting up a single *joint* hierarchical model. Partial results contribute to the particular derivation of posterior distributions without loss of information. On the other hand, a classical (frequentist) approach would require independent analyses, and carrying over results from one model to the next would result in loss of information (due to the summarizing nature of the classical (frequentist) approach).

Further details on hierarchical models can be found in [12].

3. Case Study: Measurement Campaign Goal and Setup

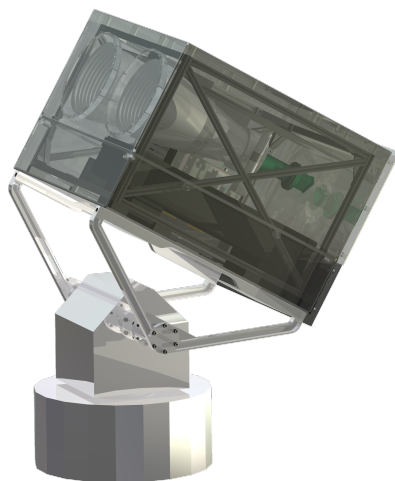
The following case study applies hierarchical Bayesian data analysis to a practical problem. This Section 3 describes the campaign goal and setup, whereas the following Section 4 applies the proposed method of hierarchical Bayesian data analysis for the estimation of a parameter in absolute radiometric calibration.

3.1. Introduction and Goal

The DLR currently develops and manufactures a set of next-generation active radar calibrators (transponders) [18], see Figure 1. These Kalibri transponders will be used in the upcoming C-band

Sentinel-1 SAR mission. The most critical transponder parameter is its radar reflectivity expressed in ERCS because the transponder will be used as a calibration standard for absolute radiometric calibration. Any error in the calibration of the calibration standards has a direct impact on the derived absolute calibration factor for Sentinel-1. First ERCS measurements of the Kalibri prototype were performed in DLR's compact antenna test range.

Figure 1. Artist's rendering of DLR's new C-band Kalibri transponder, mounted on a two-axis positioner.



The campaign was conducted for a second, independent derivation and verification of the Kalibri prototype ERCS. It consisted of eight repeated RADARSAT-2 acquisitions of a test site in which 15 trihedral reference corner reflectors of two different sizes (inner leg lengths of 1.5 m and 3.0 m) and the Kalibri transponder prototype were deployed and aligned for the respective acquisition. The data were acquired in April 2013.

The goal of the campaign was to derive the estimated transponder ERCS *including its confidence interval*. The data analysis, see Section 4, uses hierarchical Bayesian modeling as described above. The research question answered in the case study, derivation of the transponder ERCS ζ_t , is conceptually identical to the derivation of the absolute calibration factor K (linking SAR system indications P to the measurement quantity, (E)RCS) because only the roles of knowns and unknowns are reversed:

$$\zeta_t = \frac{P}{K} \quad \Longleftrightarrow \quad K = \frac{P}{\zeta_t} \quad (2)$$

Therefore, the approach used in the case study is equally applicable to the radiometric calibration of SAR systems.

3.2. RADARSAT-2 Products

The RADARSAT-2 SAR system operates at a center frequency of 5.405 GHz. The mode-dependent bandwidth goes up to 100 MHz. Center frequency and maximal bandwidth are therefore identical to those of the Sentinel-1 mission for which the Kalibri transponder was designed.

All eight RADARSAT-2 products were acquired in the *Wide Ultra-Fine* mode in VV polarization and delivered as single-look complex (SLC) images. The products offer a high nominal single-look

geometric resolution of 1.6 m × 2.8 m (range × azimuth resolution) [19]. A high-resolution mode was chosen as to increase the target-to-clutter ratio for the used point targets. Special care had to be taken in choosing the dynamic range of the processed products so that the peak of the transponder pulse response is still within the processed image's dynamic range.

An overview of all overpass times and beams is shown in Table 1.

Table 1. Overview of approximate acquisition times (in UTC).

Overpass Time	Orbit Direction	Beam Mode
7 April 2013 17:11:09	ascending	U17W2
8 April 2013 05:20:16	descending	U16W2
14 April 2013 17:06:59	ascending	U11W2
15 April 2013 05:16:06	descending	U22W2
18 April 2013 05:28:36	descending	U5W2
21 April 2013 17:02:49	ascending	U5W2
24 April 2013 17:15:19	ascending	U22W2
25 April 2013 05:24:26	descending	U10W2

3.3. Reference Point Targets

Within this study, the RADARSAT-2 system was considered radiometrically uncalibrated because of the mismatch between the system's specified relative radiometric calibration (<1 dB) [19] and the target uncertainty of about 0.2 dB. Therefore, reference point targets have been placed in the imaged scenes for manual, day-to-day calibration.

In total 15 trihedral corner reflectors were used as the reference point targets to derive the transponder ERCS. The comparatively large number was deemed necessary in order to profit from averaging during data analysis. Reflectors with two different inner leg lengths were deployed, resulting in two distinct values for their ERCS.

The RADARSAT-2 system operates at a center frequency of 5.405 GHz with a small relative bandwidth of 2%. Under this precondition, a corner's ERCS is, with sufficient accuracy, equal to a corner's RCS at the center frequency [15]. The peak RCS of an ideal corner reflector can be estimated with the physical optics approximation

$$\zeta = \frac{4\pi a^4}{3 \lambda^2} \quad (3)$$

where a is the corner's inner leg length and λ is the wavelength [20]. The resulting values are listed in Table 2.

Table 2. Corner reflectors with two different inner-leg lengths were used during the campaign.

Size	Peak RCS	Number of Targets
1.5 m	38.38 dBm ²	9
3.0 m	50.43 dBm ²	6

3.4. Target Alignment

The corner reflectors were realigned together with the transponder for each upcoming overpass, so that all corners could be used for all overpasses during analysis. The alignment angles were computed based on the predicted RADARSAT-2 orbit for the respective overpass and the point target location (latitude, longitude).

An accurate corner alignment is necessary because the peak RCS (*i.e.*, at perfect alignment) is taken as the reference value for the computation of the transponder ERCS. The corner reflector RCS is comparably unsusceptible to misalignments. On the other hand, the transponder's antenna pattern falls off more rapidly if a deviation from the main beam direction occurs, so more care needed to be taken when a transponder was aligned.

The alignment of the corner reflectors (made out of aluminum) was performed manually with an inclinometer and a compass. A local magnetic declination of 2.5° was accounted for when using a compass during azimuth alignment. The alignment standard uncertainty for both axes was estimated to be not above 0.5° .

The transponder was mounted on a positioner unit, which allows a motorized two-axis alignment in azimuth and elevation with high mechanical repeatability (better 0.1°). The alignment was zeroed in elevation with a water-level. In azimuth, a compass could not be used like for the corners because the positioner's ferromagnetic components result in a misreading. Instead, the true North direction was determined by measuring a reference azimuth direction with a GPS device. For this, two reference points were placed several dozens of meters away from the transponder, one in front, one behind, and both on the line of the current alignment. By measuring the accurate GPS positions, the azimuth orientation could be computed by determining the angle between the connecting line and true North. The standard uncertainty of this method was estimated to be 0.1° through cross-validation.

3.5. Imaged Area

All acquired RADARSAT-2 scenes imaged an area around the DLR site at Oberpfaffenhofen, Germany. The transponder was built up right on the DLR premises, allowing easy wiring and monitoring. The 1.5 m corners were installed in the immediate vicinity on the protected grasslands surrounding the airstrip at Oberpfaffenhofen. The site names for sites on which the 1.5 m corners are installed begin with *D26*. The locations are shown in Figure 2. The 3.0 m corners were located within a larger rectangular area of about $14\text{ km} \times 8\text{ km}$.

The site names for locations with a 3.0 m corner are *D24*, *D25*, and *D27* to *D30*. An overview of the 3.0 m corner sites is shown in Figure 3.

4. Case Study: Data Analysis and Results

4.1. Overview

The RADARSAT-2 datatakes were processed by MDA and the final single-look complex (SLC) images were the starting point for the data analysis.

Figure 2. Locations of the transponder on the DLR premises and of the 1.5 m corners on the adjacent airport in Oberpfaffenhofen, Germany. (Map tiles in this and Figure 3 reproduced with permission from Stamen Design, based on data from OpenStreetMap.)

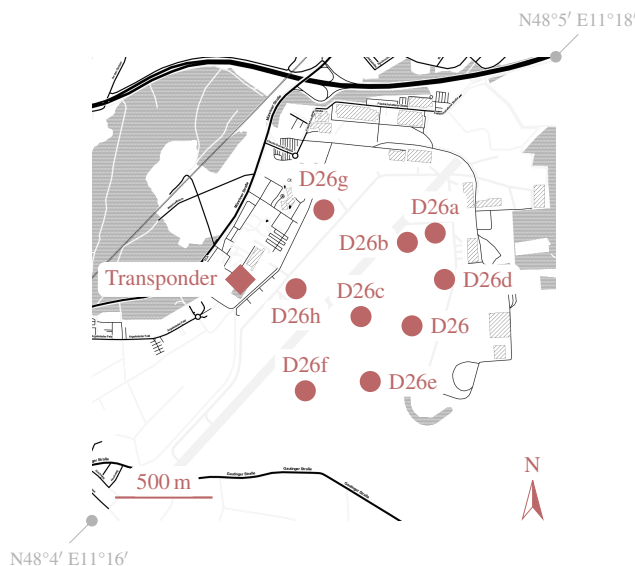
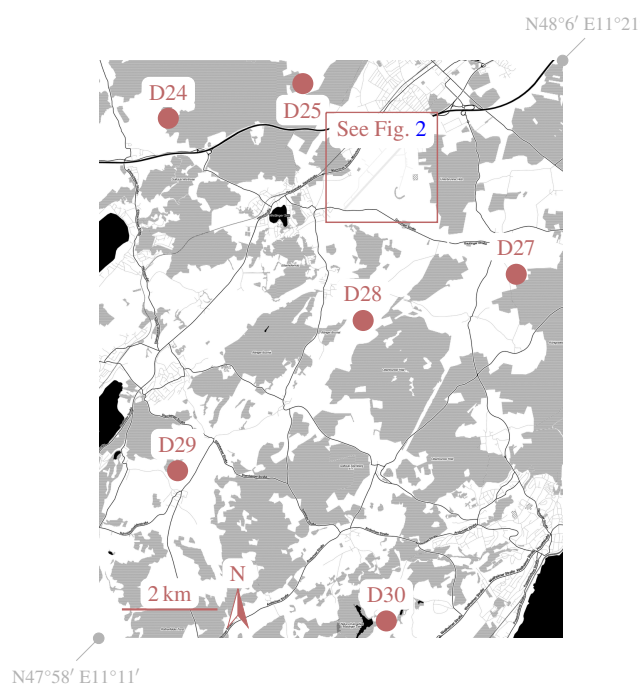


Figure 3. Map of imaged area, showing the locations of the 3.0 m corners.



The primary analysis goal is to derive the transponder ERCS and, equally importantly, an associated uncertainty statement. In principal, the transponder ERCS ζ_t can be derived if the ERCS of a reference target (placed in the same scene) ζ_r is known according to the proportionality

$$\zeta_t = \frac{P_t}{P_r} \zeta_r \tag{4}$$

Here, P_t and P_r are the point target intensities of the transponder and the reference target, derived from the processed SAR images and possibly expressed as digital numbers.

The analysis is split into two parts:

1. Point target analysis: Extract the relative point target impulse response powers for all point targets in all scenes (see Section 4.2).
2. Parameter estimation: Set up a statistical model to derive the estimated transponder ERCS and corresponding uncertainty from all datatakes (see Section 4.3).

4.2. Power Estimation for Point Targets from SAR Images

The imaged point targets appear as bright crosses in the processed image. Two methods are distinguished when deriving the point target power: the peak and the integral method. With the peak method, the target power is described by the pixel with the peak power (complex pixel amplitude squared) whereas with the integral method, the target power results as the sum of pixel powers over a relevant region around the target. It was shown that the integral method is advantageous as it leads to accurate results even if the image is not perfectly focused [21]. Thus the integral method was chosen for this analysis.

The goal of this analysis step is to derive a table of relative point target intensities for all point targets in all images. Although an absolute scale is not necessary (and the RADARSAT-2 absolute calibration factor is ignored here), it is crucial that the point target intensities derived from different images are all calibrated with respect to each other. This is achieved by applying the (range-independent) *beta naught* look-up table delivered together with the images [22]. It was confirmed that no obvious incidence-angle dependence remained in the data.

The following steps were performed to derive a targets' integrated impulse response intensity:

1. Define a search window around the point target in the georeferenced, processed image.
2. Find and record the brightest pixel location.
3. Define an analysis window, centered on the brightest pixel of the previous step.
4. Estimate the clutter power from four non-overlapping areas surrounding the peak.
5. Integrate the target power over a cross area (see Figure 4 and Table 3), capturing all pixels with significant point target power.
6. Subtract the estimated clutter power from the integrated target power to get a clutter-compensated target power.

This procedure is in line with the one described in [3].

Figure 4. Integration area for the *integral method*.

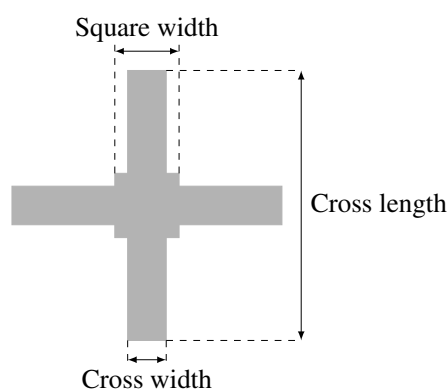
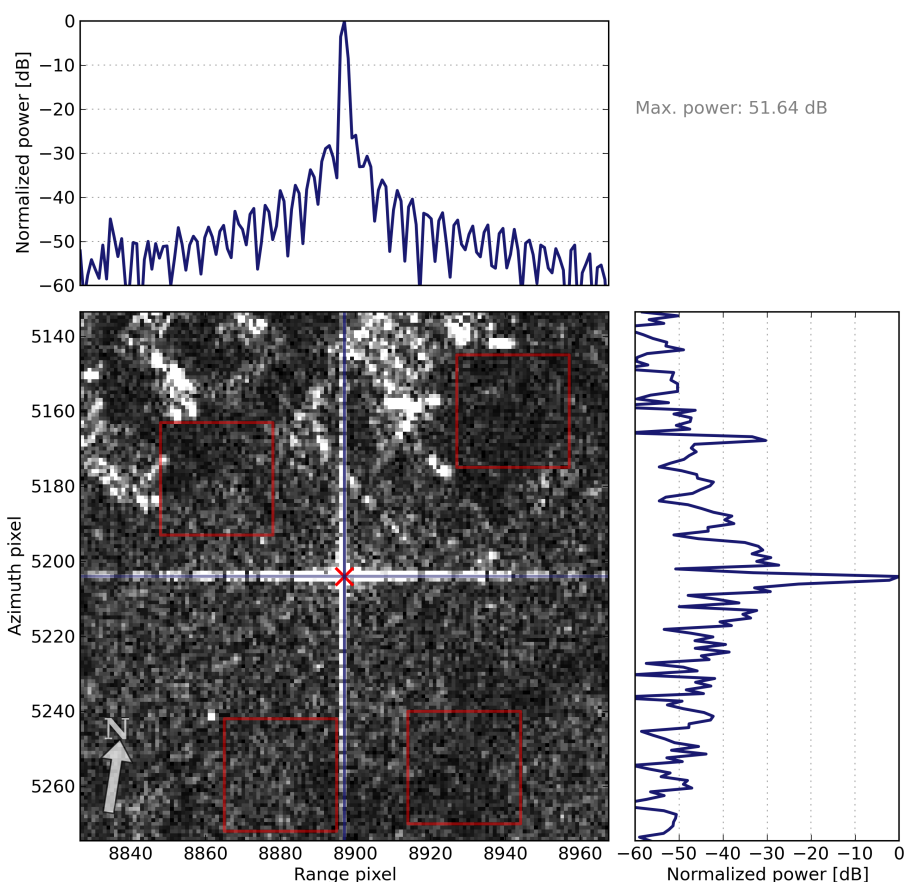


Table 3. Cross parameters according to Figure 4 used during analysis.

Parameter	Pixels
Cross length	21
Cross width	3
Square width	5

An exemplary transponder impulse response is shown in Figure 5. The peak power lies more than 60dB above the clutter level. This large separation is already an indication that clutter power compensation, in practice, is not necessary. Nevertheless, it was performed for all 126 analyzed point targets.

Figure 5. Transponder impulse response for the first overpass on 7 April 2013. A large target-to-clutter ratio is apparent. The four red squares indicate the areas from which the clutter power was estimated.



Exemplary impulse responses for 3.0 m and 1.5 m corners are shown in Figures 6 and 7, respectively. Although the impulse response is clearly visible in the range and azimuth cuts, both corners are comparatively hard to spot in the image patches themselves. This is due to the way the intensities in the image patches have been clipped for visualization. All digital values above a threshold (2.7 times the mean absolute amplitude within the patch) are set to the threshold value.

Figure 6. 3.0 corner impulse response at site D24 for the first overpass on 7 April 2013. Also see caption of Figure 5.

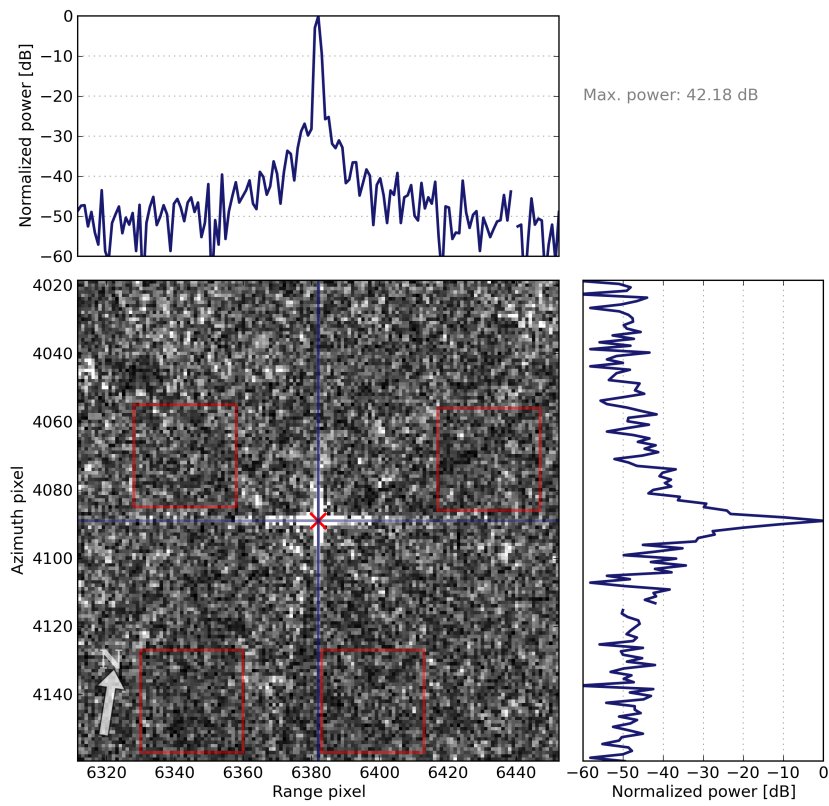
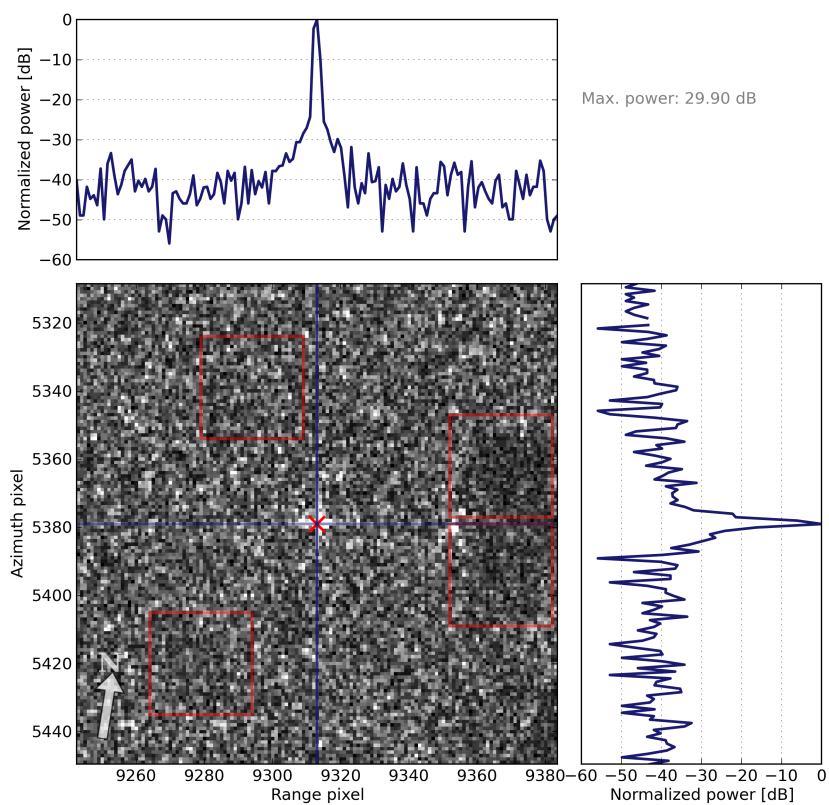


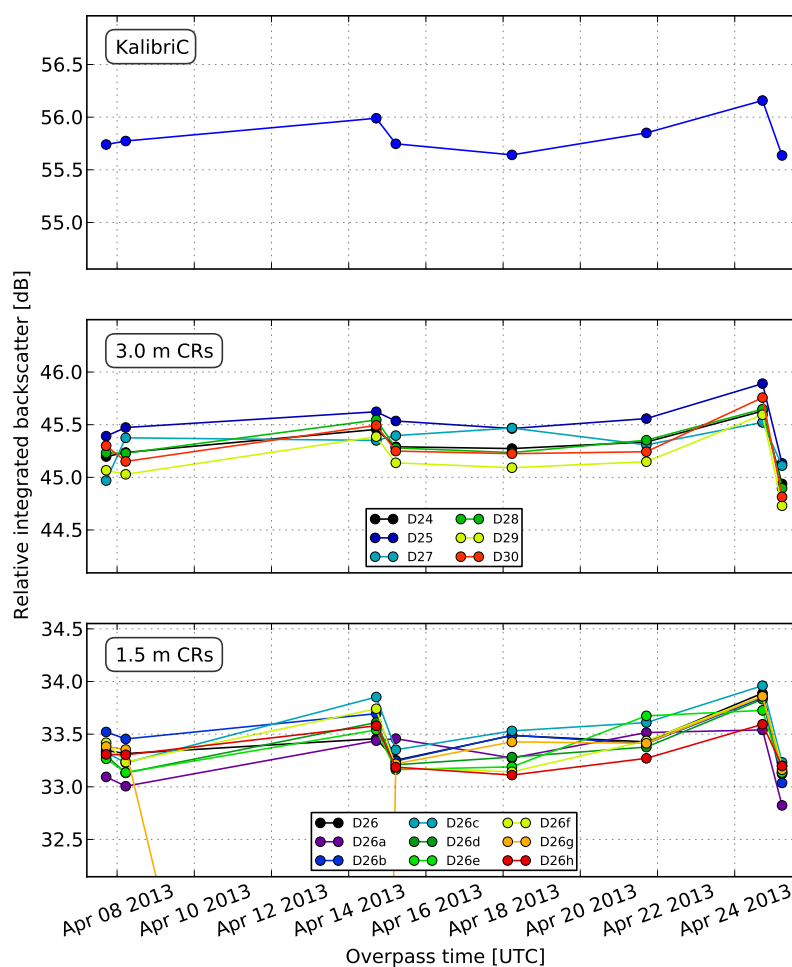
Figure 7. 1.5 corner impulse response at site D26 for the first overpass on 7 April 2013. Also see caption of Figure 5.



4.3. Bayesian Statistics and Hierarchical Model Fitting

The output of the previous processing step is a table which reports a relative point target power per target and scene. A graphical representation is shown in Figure 8. This is the input data to derive the transponder ERCS and its confidence interval.

Figure 8. Uncompensated and unmasked data, the immediate result of the processing described in Section 4.2. The data points lie on a common ordinate; for better visibility, one region per target group is plotted.



In order to estimate the target powers from all available data, daily RADARSAT-2 and transponder drifts need to be estimated and compensated, as described in the following Section 4.3.1. Then, the transponder ERCS can be computed with the measurement model Equation (4). The final, combined uncertainty shall include the uncertainties which are incorporated in each step. This problem is solved with Bayesian statistics, which directly allows to estimate the uncertainty for all derived parameters.

4.3.1. Daily RADARSAT-2 and Transponder Drifts

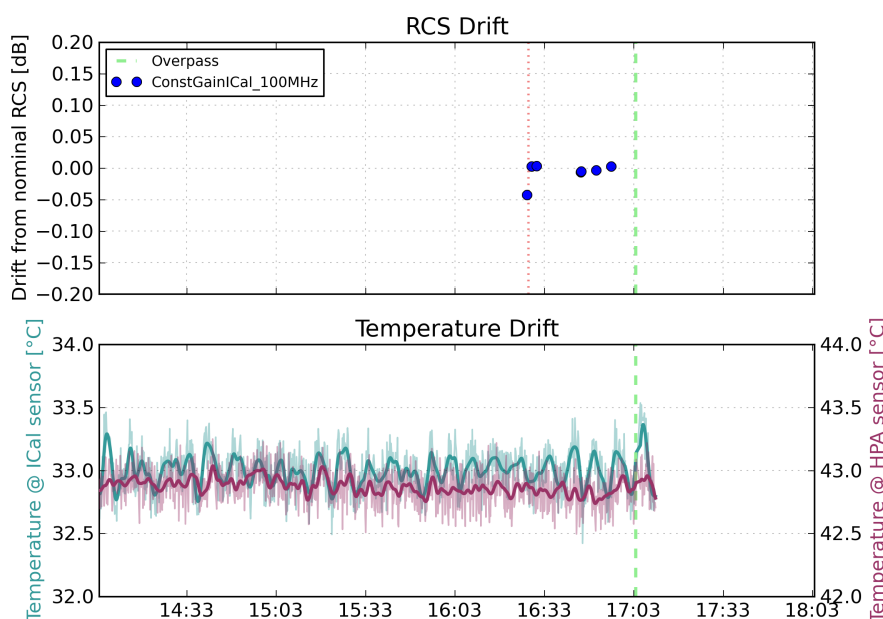
A plot of the original data from the first processing step was show in Figure 8. On first impression, the data non-uniformity is dominated by a daily systematic drift. For instance, independent of the radar target group (transponder, 3.0m, or 1.5 m corners), a radiometric drift of more than 0.5 dB can

be observed between the last and the previous overpass. Therefore, a daily RADARSAT-2 radiometric drift compensation had to be included in the analysis.

Also, there is one immediately apparent data outlier: The corner at site D26g on the third day was clearly not aligned and was masked out prior to further analysis.

Besides the RADARSAT-2 drift, the (mostly temperature-dependent) transponder drift can be estimated and compensated. The compensation is based on internal calibration data recorded by the transponder itself. Exemplary transponder loop gain and temperature drift data are shown in Figure 9. The dashed red line in the upper plot defines the point in time when the transponder RCS was corrected to its nominal RCS. The blue markers, on the other hand, indicate times at which the transponder RCS was merely estimated. In the case of Figure 9, the drift at the overpass time can be estimated to be nonexistent (0.0dB). An upper bound on the uncertainty of this estimated drift can be stated from the plot as 0.02 dB, *i.e.*, the true drift should lie, with high probability, in the range $[-0.02, 0.02]$ dB.

Figure 9. Estimated transponder drift and temperature stability (main influence for gain drift) for the overpass on 21 April 2013. The transponder operates with its nominal loop gain if a relative drift of 0dB is detected. The loop gain was adjusted at 16:30 (red dotted line), after which the drift was monitored (blue circles) until the overpass (dashed green line). After drift correction, the drift is estimated to be with high probability within the range $[-0.02, 0.02]$ dB at 17:03 UTC.



A similar assessment has been performed for all other days. The estimated transponder drifts and their estimated error bounds are listed in Table 4. A higher drift and/or error is stated for overpasses when a transponder software problem prohibited a nominal, temperature controlled operation of the prototype.

4.3.2. Hierarchical Bayesian Model

The Bayesian model requires the definition of priors, *i.e.*, probability distributions for all model parameters. These shall be defined in the following.

Table 4. Daily transponder RCS drift, estimated maximal error bounds on the estimated drift, and resulting standard uncertainties according to Equation (5).

Overpass Date	Estimated Drift μ_s (dB)	Maximal Error (dB)	Resulting σ_s (dB)
2013-04-07	0.00	0.05	0.03
2013-04-08	0.00	0.02	0.01
2013-04-14	0.02	0.03	0.02
2013-04-15	-0.01	0.03	0.02
2013-04-18	0.00	0.07	0.04
2013-04-21	0.00	0.02	0.01
2013-04-24	0.05	0.05	0.03
2013-04-25	0.02	0.03	0.02

For the eight overpass days d , eight different daily RADARSAT-2 drifts r_d need to be determined. It is estimated from Figure 8 that the daily drift, expressed as a scaling factor, will certainly be in the range 0.4 to 1.6 (*i.e.*, -4dB to 2dB). The prior can thus be written as $r_d \sim U(0.4, 1.6)$, where $U(a, b)$ describes a uniform distribution with lower and upper bounds a and b , respectively.

Besides the RADARSAT-2 drift, the transponder drift needs to be modeled. The values in column three of Table 4 define maximum error bounds, *i.e.*, a lower bound a and an upper bound b . According to [7], this information can be converted to a standard uncertainty and therefore a normal distribution, where the standard deviation σ is given as

$$\sigma = \frac{1}{\sqrt{12}}(b - a) \tag{5}$$

Hence, the daily transponder drift s_d is modeled as a normal distribution $N(\mu, \sigma)$, where μ describes its mean and σ its standard deviation: $s_d \sim N(\mu_{s,d}, \sigma_{s,d})$. The best estimate $\mu_{s,d}$ is taken from Table 4, where also the $\sigma_{s,d}$ (resulting from Equation (5)) are listed.

It is assumed that all measured point target powers of all targets within one target group g (transponder t , 3.0 “30”, and 1.5 corners “15”) belong to the same population, which can be described by a normal distribution $N(\mu_g, \sigma_g)$, having the group location (mean) μ_g and group scale (standard deviation) $\sigma_g > 0$. A normal distribution was chosen because it is symmetric (no plausible reason can be found for an asymmetric distribution), and because the distribution of the measured values results from several physical effects (satellite thermal drift, satellite and target alignment errors, target stability, clutter, *etc.*) so that the central limit theorem suggests a normal distribution as well. Each μ_g and σ_g are modeled to originate from uniform distributions: $\mu_g \sim U(10^{1.5}, 10^7)$ (*i.e.*, U(15 dB, 70 dB)) and $\sigma_g \sim U(0, 10^6)$. In a way, the σ_g are nuisance parameters which are not required to derive the transponder ERCS, but they are nevertheless necessary in order to describe the normal distributions N_g .

Now, for every overpass d , the data $y_{d,g}$ is fitted depending on its group $g \in [t, 15, 30]$:

$$\begin{aligned} y_{d,30} &\sim N(r_d \mu_{30}, \sigma_{30}) \\ y_{d,15} &\sim N(r_d \mu_{15}, \sigma_{15}) \\ y_{d,t} &\sim N(r_d s_d \mu_t, \sigma_t) \end{aligned} \tag{6}$$

This way, the daily drift parameter r_d is estimated from all available data, exploiting the fact that the drift should be equal for data across all three groups.

Estimating parameters with the model Equation (6) results in estimates for the (relative) point target powers for each group (μ_t , μ_{30} , and μ_{15}) after drift compensation.

The next step is to relate the point target powers to ERCS. For this, a reference ERCS is necessary. It was chosen to be the group ERCS of the 1.5 corners. (The not more than 7 years old 1.5 corner reflectors were, mechanically speaking, in a better shape than the more than 20 years old 3.0 corners, which show apparent deformations due to damages and their continuous exposition on a field. Mechanical deformations result in a reduction of a corners monostatic ERCS because some of the incident power is reflected away from the incident beam direction. The visual observation was confirmed by looking at the ERCS dispersion within each group: $\sigma_{15} = 0.15 \text{ dBm}^2$ and $\sigma_{30} = 0.41 \text{ dBm}^2$. These observations lead to the conclusion that the ERCS of the utilized 3.0 corners should not be used as an absolute reference, and that the 1.5 corners provide a better link to an absolute ERCS. Nevertheless, the 3.0 corners and their large ERCS helped in determining more accurately the daily RADARSAT-2 drift.)

The value of the reference ERCS is modeled as $\zeta_{15} \sim N(10^{3.838} \text{ m}^2, 10^{0.02} \text{ m}^2)$ (*i.e.*, $N(38.38 \text{ dBm}^2, 0.2 \text{ dBm}^2)$). Its location is defined by Equation (3). The standard deviation, or standard uncertainty according to [7], characterizes the state of knowledge about the reference ERCS. The statement that the ERCS of 1.5 corners can be determined with Equation (3) up to a standard uncertainty of 0.2 dBm^2 is certainly the weakest point in the argument. It is based on previous experience gained from numerical field simulations on corners of the same size at X-band, and on plausibility. (A standard uncertainty of 0.2 dBm^2 is plausible because with it the theoretical RCS difference of 3.0 and 1.5 corners can be (well) explained. The theoretical difference, according to Table 2, is 12.05 dBm^2 . The difference between the estimated mean target powers (in MCMC parameters: $\mu_{30} - \mu_{15}$) is 11.92 dB , *i.e.*, 0.13 dB away from the predicted value despite the already discussed deformation of the 3.0 corners. It is still possible though that the RCS of all corners is, due to deformation and the approximate nature of Equation (3), lower than assumed. Nevertheless, Equation (3) seems to characterize the absolute RCS of trihedral corners despite some mechanical deformations with a standard uncertainty of not more than 0.2 dBm^2 .) Nevertheless, it cannot be proofed and further work should be conducted in determining the absolute knowledge of a trihedral corner reflector's ERCS.

Now, the final transponder ERCS is deterministically related to the already derived model parameters through

$$\zeta_t = \frac{\mu_t}{\mu_{15}} \zeta_{15} \quad (7)$$

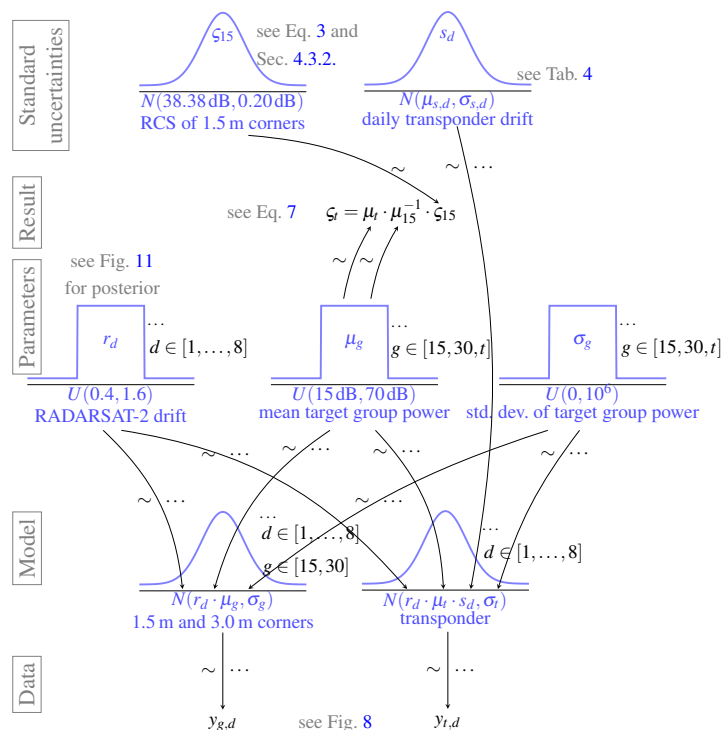
according to the measurement model Equation (4).

The complete Bayesian model described above is visualized in Figure 10.

4.3.3. Posterior Simulation

The hierarchical model developed in the previous section is now solved iteratively with the numerical Markov chain Monte Carlo method (MCMC) [12,13]. The goal is to find the most probable set of parameters (*e.g.*, r_d , μ_{30} , *etc.*) which is most likely in describing the given data.

Figure 10. Diagram of the Bayesian model. The ellipsis symbol ... indicates a family of probability distributions (per group g or overpass d), whereas \sim means that a parameter is drawn from the respective distribution.



If the model is well posed, then the iterative MCMC will converge to the true distribution of the parameters. This also means that the first draws/simulation runs need to be discarded, and only values after this *burn in* period should be considered. To improve the (required) independence between successive draws, only every n th simulation draw is considered, *i.e.*, *thinning* is applied.

In order to compute the parameters of the hierarchical model, which was presented in the previous chapter, 2×10^5 simulation runs were conducted, allowing for a burn in of 1×10^4 and a thinning of 20. These parameters were determined empirically by observing the traces and autocorrelations of the model parameter draws.

4.3.4. MCMC Results

The solution of the hierarchical model describes all parameters at the same time. Nevertheless, the results can be visualized step by step.

The first result is the estimated RADARSAT-2 drift. The estimated drift is shown in Figure 11. The error bars indicate the range of values which defines the 95% probability intervals. From this it can be seen that the drift between the first and the second overpass is statistically not significant, but the drift between the last two overpasses, for instance, is. The drift was estimated with a standard uncertainty (according to [7]) of always less than 0.1 dB.

This estimated RADARSAT-2 drift can now be applied to the measured data. The original data in Figure 8 appears now much more uniform, see Figure 12. No apparent systematic drift remains.

Figure 11. Estimated daily drifts r_d of the RADARSAT-2 system. The error bars indicate 95% highest-probability density intervals.

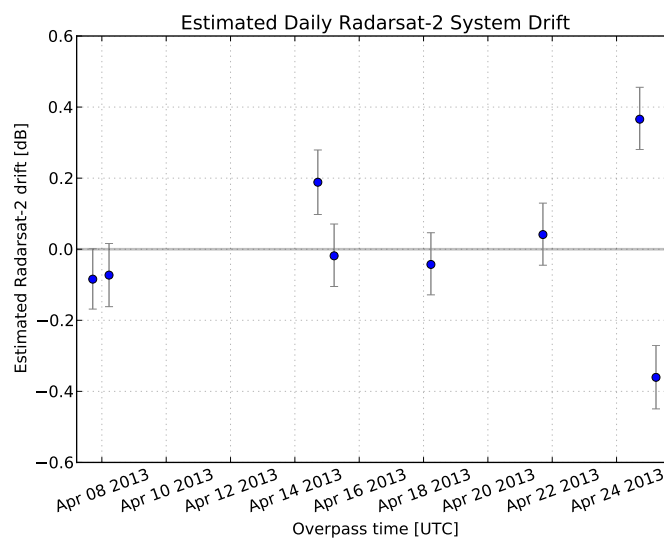
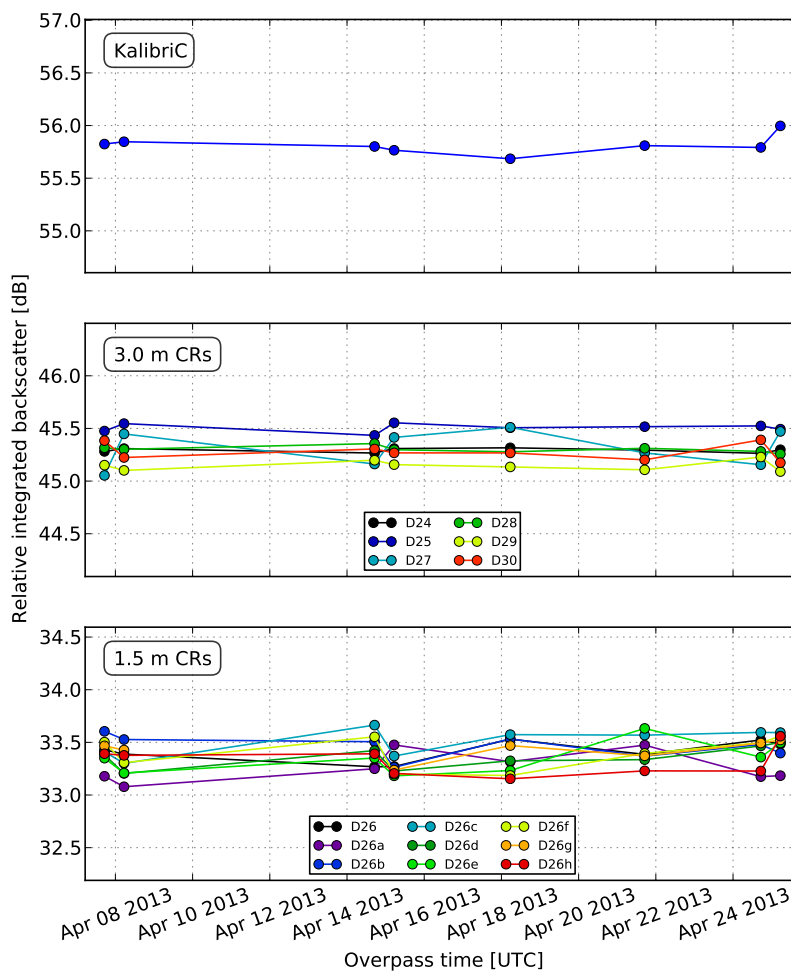
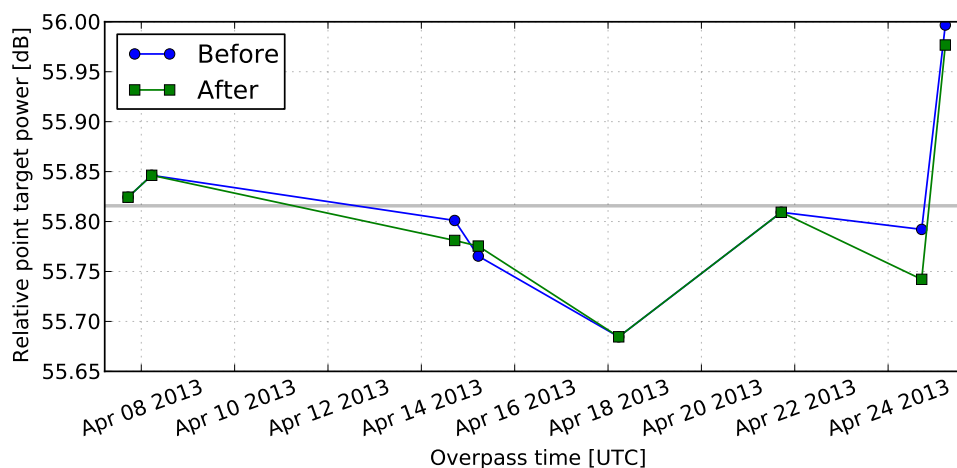


Figure 12. Measured impulse response powers with RADARSAT-2 drift compensation (see Figure 11) applied. The dispersion within one target group is now greatly reduced with respect to Figure 8.



After RADARSAT-2 drift compensation, the estimated transponder drift can be taken out from the upper plot in Figure 12. The plot in Figure 13 results. Already visually it becomes clear that the transponder drift is small in comparison to the remaining dispersion of the measurement data. This can also explain why at times the drift compensation results in updated values which lie further away from the population mean.

Figure 13. Visualization of the Kalibri transponder drift compensation with data from Table 4. For visual guidance, the gray line marks the sample mean of the data before transponder drift compensation.



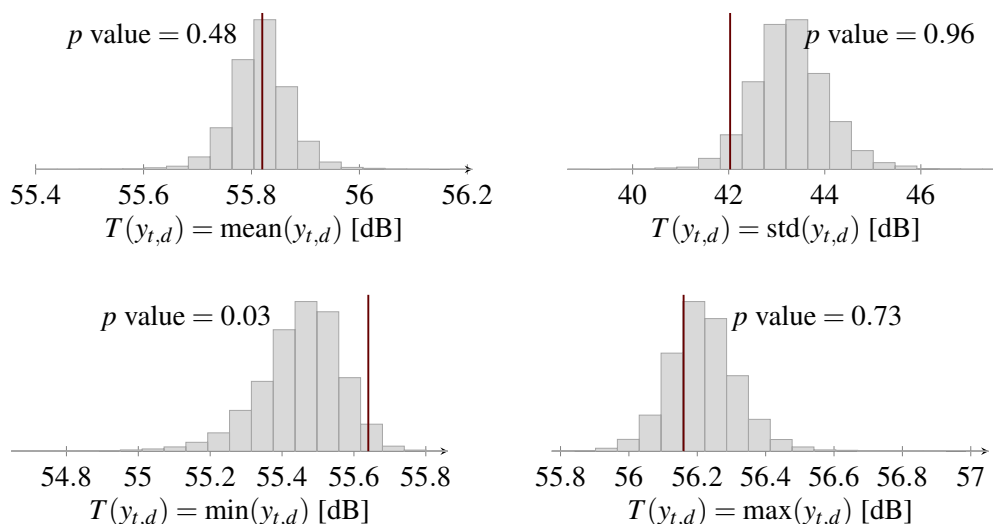
Now the most important result of the MCMC simulation is the derivation of the transponder ERCS ζ_t and its standard uncertainty. The transponder ERCS was estimated to be 60.80 dBm² with a standard uncertainty according to [7] of 0.206 dBm². The 95% highest probability density interval is given as [60.38, 61.17] dBm². Note that the standard uncertainty is clearly dominated by the assumed ERCS knowledge uncertainty of the 1.5 corners of 0.2 dBm².

4.4. Posterior Predictive Checks: Model Verification

In Section 4.3.2, a normal distribution was chosen in order to model the observed integrated pixel intensities. Here it shall be demonstrated that this model is indeed plausible and adequate.

Focusing on the transponder data $y_{t,d}$, test statistics $T(y_{t,d})$ of the observed data can be compared to the test statistics of replicated data samples $T(y_{t,d}^{rep})$, i.e., samples which were generated numerically through the model [12]. If this analysis is conducted for the four test statistics mean, standard deviation, minimum, and maximum value across all eight overpasses, Figure 14 results. A good model fit is found if the test statistic of the observed data (vertical line) lies close to the center of mass of the histogram. As a criterion, the p value (relative number of samples above observed test statistic) can be used. At a confidence level of 95%, the p value should therefore be within the range of 2.5% to 97.5%. This is observed for all four test statistics, and especially the most important aspect of the model, its mean, is well reproduced by the model with a p value close to 1/2.

Figure 14. Posterior predictive checking for predicted (modeled) transponder data $y_{t,d}^{\text{rep}}$ and four different test statistics. Especially the most important aspect of the predicted data, its mean, is well modeled.



4.5. Plausibility Check with Classical Statistics

As a means of verification, the result of the previous section can be reproduced approximately with classical (frequentist) statistics. One way of handling the hierarchical data structure is to derive one transponder ERCS per overpass, and then to combine the resulting eight ERCS values through averaging into an overall transponder ERCS. The disadvantage of this simplified approach (in comparison to the approach using hierarchical Bayesian data analysis as shown before) is that information about the uncertainty of each of the eight transponder ERCS values is lost and does not contribute to the combined uncertainty.

After averaging, an estimated standard deviation of the mean can be derived from the eight estimated ERCS values, resulting in a standard uncertainty for the estimated transponder ERCS [7]. The third step is then to derive a combined standard uncertainty by incorporating the ERCS knowledge uncertainty of the 1.5 corners of 0.2 dB through the method of “root-sum-square”. Lastly, the expanded uncertainty [7] with a coverage factor of $k = 2$ is derived.

This approach results in an estimated transponder ERCS of 60.80 dBm² with a standard uncertainty of 0.20 dB or an expanded standard uncertainty of 0.41 dB at a confidence level of 95% ($k = 2$). The result confirms the findings of the previous section.

Note that once again the combined uncertainty is dominated by the assumed ERCS knowledge uncertainty of the 1.5 corner reflectors. The transponder’s combined backscatter uncertainty is sufficiently low to permit the calibration of SAR instruments like Sentinel-1 to an absolute radiometric uncertainty of 1.00 dB at a confidence level of 99%, provided that the SAR instrument is otherwise sufficiently precise [23].

5. Discussion of Hierarchical Bayesian Data Analysis for Radiometric Calibration

The advantages of using hierarchical Bayesian data analysis for radiometric calibration were laid out before: The approach can jointly answer typical analysis questions in radiometric calibration while fully exploiting the hierarchical nature of external calibration data and fulfilling the requirement on reporting measurement uncertainties and confidence intervals. This section shall add a critical discussion of the proposed method.

First, the most recognized approach in metrology for deriving measurement uncertainties is the ISO *Guide to the expression of uncertainty in measurements* (GUM) [7]. The GUM, which was introduced in 1993, fundamentally applies classical (frequentist) statistics, and is not directly compatible with a Bayesian approach (which was proposed in this paper). Nevertheless, the current GUM approach has been repeatedly criticized, and Bayesian methods have been proposed as a consistent and universally applicable replacement [24–26]. In practice, uncertainties derived by Bayesian statistics are often equal or approximately equal to uncertainties derived by classical statistics [25] so that both approaches lead to desired results. Yet thanks to the availability of numerical tools, Bayesian computations are now often simpler than classical approaches if hierarchical data is analyzed [14]. Therefore, taking a Bayesian approach for deriving the measurement uncertainty of the absolute calibration factor seems justified.

From the outset, using Markov chain Monte Carlo (MCMC) simulations to infer model parameters appears to be more complicated than employing classical statistics. The added initial work is offset though by a joint probability model, which allows to derive model parameters on arbitrary hierarchical levels without loss of information. Hence the improved analysis justifies the initial extra work.

If numerical methods like MCMC are used, problems of non-convergence can occur and must be addressed during analysis. Care must be taken in the assessment of simulation results, and plausibility checks should be added.

6. Conclusions

The presented work proposed for the first time in the field of radiometric SAR system calibration to exploit hierarchical Bayesian data analysis. It was claimed that Bayesian statistics is well suited to the analysis of calibration data because of the following key factors:

- Within Bayesian statistics, probability distributions are used in describing model parameters. The distributions convey a meaning of uncertainty. Bayesian statistics is therefore an appropriate choice for calibration, where an estimated parameter is meaningless without a statement of its uncertainty.
- Hierarchical joint probability models are well suited to describe data that is typically acquired during an external radiometric SAR calibration campaign. During data analysis, depending on the research question, parameters often need to be estimated on different levels or for different groups. Hierarchical Bayesian modeling is well suited to derive model parameters for different interdependent parameters, especially when numerical methods like Markov chain Monte Carlo simulations are used.

The applicability of the method for radiometric calibration was demonstrated through a case study. The data of an external calibration campaign was analyzed. The campaign goal was to derive the ERCS

of DLR's C-band Kalibri transponder prototype. Due to hierarchical Bayesian data analysis, it was possible to estimate and compensate the overpass-dependent drift of the RADARSAT-2 system and to derive the transponder ERCS with a remaining standard uncertainty (according to GUM) of only 0.21 dBm².

In order to convert the case study approach to an operational uncertainty analysis procedure for SAR missions, a database of point-target measurements should be set up. Filling the database incrementally with measurements of permanently installed reference point targets over the mission lifetime would allow one to continually derive radiometric uncertainty estimates based on Bayesian statistics.

The authors are convinced that Bayesian statistics and hierarchical modeling are an important step toward *traceable* radiometric SAR system calibration. Uncertainties and confidence intervals can now be derived with the same diligence that is presently applied to the estimation of calibration parameters.

Acknowledgments

The campaign was only feasible due to the commitment and collaboration of several partners and many DLR colleagues. Ron Caves from MDA Systems Ltd. rendered the campaign possible by organizing the acquisitions and by providing the processed RADARSAT-2 datatakes. The EDMO airport in Oberpfaffenhofen and namely Werner Döhring greatly simplified the campaign work by permitting the installation of all utilized 1.5 corners right on the grassland surrounding the airstrip, directly next to the DLR premises.

On the side of DLR, the campaign was made possible by Manfred Zink and the additional campaign team members Christo Grigorov, Thomas Kraus, and Sebastian Ruess.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Van Zyl, J.J.; Kim, Y. *Synthetic Aperture Radar Polarimetry*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2011.
2. Dobson, M.C.; Ulaby, F.T.; Letoan, T.; Beaudoin, A.; Kasischke, E.S.; Christensen, N. Dependence of radar backscatter on coniferous forest biomass. *IEEE Trans. Geosci. Remote Sens. Electr.* **1992**, *30*, 412–415.
3. Freeman, A. SAR calibration: An overview. *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 1107–1121.
4. Curlander, J.C. *Synthetic Aperture Radar: Systems and Signal Processing*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1991.
5. Schwerdt, M.; Bräutigam, B.; Bachmann, M.; Döring, B.; Schrank, D.; Hueso Gonzalez, J. Final TerraSAR-X calibration results based on novel efficient methods. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 677–689.
6. Shimada, M.; Isoguchi, O.; Tadono, T.; Isono, K. PALSAR radiometric and geometric calibration. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 3915–3932.

7. *Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement*; ISO/IEC Guide 98-3:2008; ISO copyright office: Geneva, Switzerland, 2008.
8. Schwerdt, M.; Hueso Gonzalez, J.; Bachmann, M.; Schrank, D.; Döring, B.; Tous Ramon, N.; Walter Antony, J.M. In-Orbit Calibration of the TanDEM-X System. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Vancouver, BC, Canada, 24–29 July 2011; pp. 2420–2423.
9. Luscombe, A.P.; Thompson, A.A. RADARSAT-2 Calibration: Proposed Targets and Techniques. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Vancouver, BC, Canada, 24–29 July 2001; pp. 496–498.
10. Luscombe, A. Image Quality and Calibration of RADARSAT-2. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Cape Town, South Africa, 12–17 July 2009; pp. 757–760.
11. Chakraborty, M.; Panigrahy, S.; Rajawat, A.S.; Kumar, R.; Murthy, T.V.R.; Haldar, D.; Chakraborty, A.; Kumar, T.; Rode, S.; Kumar, H.; *et al.* Initial results using RISAT-1 C-band SAR data. *Curr. Sci.* **2013**, *104*, 490.
12. Gelman, A.; Carlin, J.B.; Stern, H.S.; Rubin, D.B. *Bayesian Data Analysis*, 2nd ed.; Chapman & Hall/CRC: Boca Raton, FL, USA, 2004.
13. Gregory, P. *Bayesian Logical Data Analysis for the Physical Sciences*; Cambridge University Press: Cambridge, UK, 2005.
14. Bayarri, M.J.; Berger, J.O. The interplay of bayesian and frequentist analysis. *Stat. Sci.* **2004**, *19*, 58–80.
15. Döring, B.J.; Schwerdt, M. The radiometric measurement quantity for SAR images. *IEEE Trans. Geosci. Remote Sens.* **2013**, doi: 10.1109/TGRS.2012.2234128.
16. Bolstad, W.M. *Introduction to Bayesian Statistics*, 2nd ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2007; p. 464.
17. Patil, A.; Huard, D.; Fonnesbeck, C.J. PyMC: Bayesian stochastic modelling in Python. *J. Stat. Softw.* **2010**, *35*, 1–81.
18. Döring, B.J.; Looser, P.; Jirousek, M.; Schwerdt, M.; Peichl, M. Highly Accurate Calibration Target for Multiple Mode SAR Systems. In Proceedings of the European Conference on Synthetic Aperture Radar, Aachen, Germany, 7–10 June 2010; Volume 8.
19. MDA Corporation. *Radarsat-2 Product Description*; MacDonald, Dettwiler and Associates Ltd: Richmond, BC, Canada, 2011.
20. Oerry, A.W.; Brock, B.C. *Radar Cross Section of Triangular Trihedral Reflector with Extended Bottom Plate*; Technical Report May; Sandia National Laboratories: Albuquerque, NM, USA, 2009.
21. Gray, A.L.; Vachon, P.W.; Livingstone, C.E.; Lukowski, T.I. Synthetic aperture radar calibration using reference reflectors. *IEEE Trans. Geosci. Remote Sens.* **1990**, *28*, 374–383.
22. Dettwiler, M. *RADARSAT-2 Product Format Definition*; Technical Report 1/9; MacDonald, Dettwiler and Associates Ltd.: Richmond, Canada, 2011.

23. Schwerdt, M.; Döring, B.; Zink, M.; Schrank, D. In-Orbit Calibration Plan of Sentinel-1. In Proceedings of the European Conference on Synthetic Aperture Radar, Aachen, Germany, 7–10 June 2010; pp. 350–353.
24. Weise, K.; Wöger, W. A Bayesian theory of measurement uncertainty. *Meas. Sci. Technol.* **1992**, doi: 10.1088/0957-0233/4/1/001.
25. Kacker, R.; Jones, A. On use of Bayesian statistics to make the Guide to the Expression of Uncertainty in Measurement consistent. *Metrologia* **2003**, doi: 10.1088/0026-1394/40/5/305.
26. Willink, R.; White, R. *Disentangling Classical and Bayesian Approaches to Uncertainty Analysis*; Technical Report No. CCT/12-08; BIPM: Sèvres, France, 2012.

© 2013 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).