Monitoring the Absolute Calibration of a Polarimetric Weather Radar

MICHAEL FRECH
Meteorologisches Observatorium Hohenpeissenberg, Deutscher Wetterdienst, Hohenpeissenberg, Germany

MARTIN HAGEN
Institut für Physik der Atmosphäre, DLR, Oberpfaffenhofen, Germany

THEO MAMMEN
Deutscher Wetterdienst, Hamburg, Germany

(Manuscript received 6 April 2016, in final form 14 December 2016)

ABSTRACT

The absolute calibration of a dual-polarization radar of the German Weather Service is continuously monitored using the operational birdbath scan and collocated disdrometer measurements at the Hohenpeissenberg observatory. The goal is to measure the radar reflectivity constant $Z$ better than $\pm 1$ dB. The assumption is that a disdrometer measurement close to the surface can be related to the radar measurement at the first far-field range bin. This is verified using a Micro Rain Radar (MRR). The MRR data fill the gap between the measurement near the surface and the far-field range bin at 650 m. Using data from the first half of the warm season in 2014, a bias in radar calibration of 1.8 dB is found. Data from only stratiform precipitation events are considered. After adjusting the radar calibration and using an independent data sample, very good agreement is found between the radar, the MRR, and the disdrometer with a bias in $Z_{\text{hradar}}$ smaller than 1 dB. The bias in $Z_{\text{hradar}}$ is not captured with the classic one-point calibration, which is performed twice a day using a built-in test signal generator. This is attributed to the fact that the characterization of the transmit and receive path is not accurate enough. Solar interferences during the operational scanning are used to characterize the receiver. There, the bias found is small, about 0.2 dB, so that bias based on the comparison of the radar with external sensors is attributed to the transmit path. The representativeness of the disdrometer measurements are assessed using two additional disdrometers located within 200-m distance.

1. Introduction

Dual-polarization systems are now commonly introduced in operational weather radar networks all over the world. Often, the introduction of dual-polarization systems is motivated by the expected benefit of improved algorithms for quantitative precipitation estimation and better classification of meteorological and nonmeteorological targets. A crucial prerequisite for the success of algorithms is the quality of the input radar moments. For systems that operate in a simultaneous transmit and receive mode, the key moments are the horizontal radar reflectivity factor $Z_{\text{hradar}}$; the differential reflectivity $Z_{\text{DR}}$; the cross-correlation coefficient $\rho_{hv}$; the differential phase $\Phi_{dp}$; and based on latter moment, the specific differential phase $K_{dp}$. The goal is to achieve an accuracy better than 1 dB in $Z_{\text{hradar}}$ and 0.1 dB in $Z_{\text{DR}}$. In this study we focus on monitoring the absolute calibration of $Z_{h}$. There are various methods for calibrating $Z_{h}$ (Atlas 2002; Chandrasekar et al. 2015). The standard procedure in operational networks is an engineering approach, where a HF test signal is injected into the receive path in order relate the received power to the radar reflectivity factor using the radar equation (e.g., Doviak and Zrnić 2006). This requires a good radar hardware characterization of the transmit and receive path such as, for example, the transmit power or losses. For modern radar systems, maintenance is normally scheduled every other month, so the continuous remote monitoring of the radar calibration becomes important. Such monitoring includes the analysis of operational radar data, built-in test equipment (BITE) data, and dedicated scans to extract and diagnose the system...
health. The continuous automated monitoring is essential because a problem in the system often manifests itself in a trend or subtle changes in radar data. For the accuracy of $Z_{h_{\text{radar}}}$, the goal of 1 dB is one aspect, but the other aspect is to guarantee a homogeneous quality of $Z_{h_{\text{radar}}}$ throughout the whole radar network. From our experience, this is difficult to achieve with an engineering approach because the uncertainties in the characterization of the radar hardware can vary substantially from system to system for reasons that are sometimes difficult to isolate. For example, the antenna gain is difficult to quantify within 0.5-dB absolute accuracy (e.g., Skolnik 2001).

Absolute calibration can also be determined using targets with a known radar cross section, like a metal sphere that is lifted, for example, by a balloon. However, this is not a very practical method for a radar network to monitor continuously the absolute calibration. An elegant method is the self-consistency approach to compute the bias of absolute calibration (e.g., Gorgucci et al. 1992; Ryzhkov et al. 2005; Gourley et al. 2009). From the observed and theoretical $\Phi_{dp}$, the expected $Z_h$ can be compared to the measured $Z_h$, and as such the calibration bias can be quantified. This method makes assumptions about the existing drop size distribution (DSD) but appears to be relatively insensitive to variations in the DSD but sensitive to the drop-size shape assumptions. Another recent approach for quantifying the bias in absolute calibration makes use of the specific attenuation of $Z_h$ using the self-consistency assumption (Diederich et al. 2015). Those methods appear to work quite well during the warm season, while it is not clear how they work as a year-round method for monitoring and maintaining the absolute calibration.

In this work we investigate a more classic approach for monitoring absolute calibration using a direct comparison of radar data and disdrometer measurements (e.g., Joss et al. 1968). Initially, we have started to install disdrometers at some radar sites as present weather sensors in order to test their capability for providing information about radome attenuation in case the radome surface becomes wet. Furthermore, this setup is also expected to be helpful in tracing and monitoring the degradation of the hydrophobic coating of the radome surface. As it has been shown before, that vertically pointing radars and collocated disdrometer measurements agree quite well (Gage et al. 2000; Bringi et al. 2013; Frech 2013), we looked into the potential of using the on-site disdrometer data to monitor the absolute calibration of $Z_{h_{\text{radar}}}$ and also $Z_{v_{\text{radar}}}$, the radar reflectivity of the vertical polarization. This methodology benefits from the fact that we operationally run a birdbath scan every 5 min at 90° elevation in order to calibrate and monitor $Z_{DR}$. The birdbath scan also provides a detailed look into the precipitation characteristics above the radar site, which is expected to have good potential for operational applications (Frech and Steinert 2015). Using radar data and collocated disdrometer measurements from three different radar sites showed promising results in assessing the absolute calibration of a radar system (Frech 2013). There, we compare the first far-field range bin at 650 m above the site with disdrometer data. The radar data are filtered for situations with stratiform precipitation without any brightband effects. Based on these results, it appears that the vertical variability of $Z_{h_{\text{radar}}}$, and therefore the variations in DSD must be small under these conditions. To verify this conclusion, we have installed a vertically looking Micro Rain Radar (MRR; Peters et al. 2005) that is able to fill the gap between the far field of the operational C-band radar and the disdrometer. Our setup, especially with respect to the spatial separation of the sensors, is similar to the set of the work of Tokay et al. (2009). Compared to our study, they use a vertically pointing S-band radar and the primary focus is the assessment of the agreement between the disdrometer and the remote sensing instruments, in particular with respect to the derived reflectivity–rainfall ($Z$–$R$) relationships. Calibration issues in relation to the remote sensing sensors are not discussed in their study.

The comparison of radar measurements with disdrometers or other ground-based sensors and the associated errors are addressed in various studies (e.g., Gebremichael and Krajewski 2004; Lee et al. 2009; Bringi et al. 2011; Thurai et al. 2012). The error budget is analyzed in detail for a radar pixel and a collocated disdrometer measurement 15 km away from the radar in Thurai et al. (2012). Contributions to the error budget stem from the point-to-volume variance, radar measurement error, retrieval error, and disdrometer sampling error. The magnitude of those errors is dependent on the retrieved parameter considered.

This paper is structured as follows. We first introduce the technical details of the DWD radar network and the radar systems. We introduce the standard calibration procedure of the polarimetric radar system. We then discuss the setup of the instruments at the Hohenpeissenberg observatory, where the DWD research radar is operated. We then introduce the data analysis and discuss the results and implications. In the appendix we assess the consistency of three disdrometers [two Thies disdrometers and one upgraded laser-optical OTT Particle Size and Velocity (Parsivel) disdrometer] that are separated less than 200 m. From this we infer aspects related to the spatial and temporal variability of the precipitation events. We conclude with a summary and an outlook on further developments.
2. DWD radar network and system

The German Meteorological Service (DWD) is operating the national weather radar network with 17 operational systems. Apart from one system, 16 systems have dual-polarization capabilities following a replacement project that was finished in 2015. The remaining system will also be replaced once a new site is secured. An additional system serves as a research radar that is operated at the Hohenpeissenberg observatory. There, new technologies, radar data processing algorithms, radar software, and new products are tested and evaluated before they are introduced into operational service. The radar network will consist of one system type, Enterprise Electronics Cooperation’s (EEC) Doppler weather radar DWSR5001C/SDP/CE [simultaneous dual polarization (SDP)].

We summarize briefly some technical aspects of the radar system:

Pedestal unit: Pointing accuracy of $<0.05^\circ$ and maximum azimuth rate of $4^\circ$ s$^{-1}$.

Transmitter: Magnetron-based transmitter with a peak power of 500 kW [so, 250 kW for both the horizontal (H) and vertical (V) channels]. The transmitter is operated in a frequency range of 5600–5650 MHz (C band). Four pulse widths are available: 0.4, 0.8, 2, and 3 $\mu$s, but operationally 0.4- and 0.8-$\mu$s pulse lengths are used.

Receiver: The receiver is mounted behind the antenna (‘‘receiver over elevation’’ concept). The analog signals are digitized by the ENIGMA3p intermediate frequency digitizer (IFD) and the digitized in-phase and quadrature phase (IQ) data are transmitted in real time through a fiber-optic rotary joint to the ENIGMA3p signal processor that is mounted in the radar control cabinet. The dynamic range of the dual IF receiver is $\geq105$ dB.


Antenna: The parabolic antenna by Seavey has a diameter of 4.27 m and consists of nine elements. It is made of a composite material and has a center-fed antenna design with four struts supporting the dual-polarization feed (Frech et al. 2013).
Radome: The radome has been manufactured by Antenna for Communications (AFC) and has a random panel design that is optimized for dual-polarization applications. The panels have a sandwich foam core design. The radome is coated with a highly hydrophobic material (Frech et al. 2013). The components, setup, interfaces, and data flow of the DWD radar system are shown in Fig. 1. The basic radar moments are computed in real time in the signal processor. The configuration of the signal processing and the scanning is handled by the radar operations software MURAN on the on-site radar computer. MURAN also computes standard radar products, and the software includes maintenance utilities for maintenance. In addition, on-site monitoring tools check the long-term system health and data quality (Frech 2013).

Once a full sweep has been acquired (single sweep, multimoments), it is transmitted to the central unit in Offenbach, Germany, where the central processing suite Polarimetric Radar Algorithms (POLARA) is hosted (Tracksdorf et al. 2013; Frech and Steinert 2015). Typically, 36 radar moments are transmitted with every sweep. In POLARA, quality-controlled data are used to compute quantitative precipitation estimates (QPE) and hydrometeor classification (HMC) products for all radar sites. Single-site products and composites are then visualized in the Ninjo display system.

The scan theorem comprises a precipitation scan (surveillance scan with variable elevation angle depending on orography), a volume scan (10 elevations between 0.5° and 25°), and a birdbath scan at 90° elevation, which is repeated every 5 min. In addition, a clutter target scan (illumination of a well-defined clutter target with standing antenna) is run every hour and a calibration scan is run twice a day. The calibration scan performs a one-point calibration using the built-in internal test signal generator. The results from this scan are used only for monitoring purposes. The clutter target scan is primarily used to monitor the coherence (and as such the clutter suppression capability) of the radar system.

The radar systems are maintained by DWD technicians every 9 months. During maintenance an engineering calibration is done. This involves in total four calibrations (H and V for 0.4- and 0.8-μs pulses), a check of the pulse form and the transmit power. For the calibration of the radar system, a carefully calibrated external test signal generator

Fig. 2. Calibration diagram of the DWSR5001C/SDP/CE system, illustrating the principal components of the transmit and receive parts that are needed to characterized properly for calibration.
(TSG) is used instead of the internal TSG located in the receiver box. The test signal is injected through couplers into the receive path, which are located in the receiver box. This is done after the transmit path has been characterized, which includes a measurement of the transmit peak power. The calibration diagram (Fig. 2) also schematically shows all elements in the transmit and receive path, which are characterized directly or indirectly as losses. Losses in the system, as sketched in Fig. 2, are usually not remeasured during maintenance and are assumed constant in the calibration procedure. Figure 2 illustrates the relevant components in the system that can/will have an influence on the analog signal and the reference measurement points that are used in the power calibration procedure.

3. Instrumentation setup and data analysis

In close proximity to the Hohenpeissenberg C-band radar (within less than 200-m distance), we operate an MRR (Peters et al. 2005) and two Thies Laser Precipitation Monitors [LPM; Laser-Niederschlags-Monitor (LNM); de Moraes Frasson et al. 2011] and one Parsivel disdrometer (Tokay et al. 2014; Fig. 3). For the latter instruments, a laser light sheet is illuminating a sensing volume. If precipitation is falling through the sensing volume, then the degree of attenuation relates to the particle size, and the duration of the attenuated signal links to the fall speed of the particle. Precipitation type is determined by DSD and the observed fall speeds. The radar reflectivity factor is computed using the measured DSD as follows:

$$Z_{\text{dis}} = 10 \log_{10} \sum_{i=1}^{n} D_i^6 N(D_i) \Delta D_i$$  \hspace{1cm} (1)$$

with the number of drop size classes $n$, the drop size number of classes $i$, the number concentration $N_i$, and the drop diameter $D_i$ of the respective class $i$. For both disdrometer types, we are using $Z$ as calculated by the respective firmwares.

Available are also sonic anemometer wind measurements for both disdrometers, which are used to filter out strong wind situations, where drop size spectra are likely to be biased (see, e.g., Friedrich et al. 2013). For the comparison with the radar data, we decided to use 1-min averaged reflectivity values.

When relating the in situ measurements with the radar measurement, we have to assume that the precipitation characteristics from a disdrometer are comparable to the ones observed by the radar sensing volume. We essentially make an ergodicity assumption that the time-average-based $Z$ from the disdrometer is the same as the $Z$ computed from the pulse volume and from a number of pulses. Schematically, the setup is illustrated in Fig. 4. We evaluate the first range bin in the antenna far field of the birdbath scan, that is, 650 m above the radar site. The assumption is that the DSD and hydrometeor type does not change significantly between this height level and the surface, where the disdrometer is situated. We also require that the brightband bottom is above 650 m. In the following analysis, we use data from the disdrometer d1. The pulse volume of the MRR at 650-m height is about $11,350 \text{ m}^3$ ($2^\circ$ beamwidth and 50-m range resolution), and for the radar it is about $18,850 \text{ m}^3$ (beamwidth of $1^\circ$ and range resolution pulse length of 0.8 $\mu$s).
The basic settings of the birdbath scan are as follows:

- **Elevation**: 90°
- **Azimuth speed**: 48° s⁻¹
- **Dynamic angle syncing (DAS)**: 5°
- **Range sampling**: 25 m
- **PRF**: 2400/1500 Hz
- **Pulse width**: 0.4 and 0.8 μs
- **Range**: 20 km

The scan is specified such that it can run in a very short period. We therefore use the maximum antenna rotation speed. Since we do not need any azimuth resolution at this elevation, we specify the largest possible DAS interval of 5° to sample as many pulses as possible within a ray. At a given range, radar moments are averaged over all azimuth angles; thus, the canting effects on $Z_{\text{DR}}$ are averaged out.

The MRR data are compared to the first far-field range bin of the C-band radar and to assess the variability of $Z_{\text{radar}}$ with height using the MRR data, closing the gap between the disdrometer and radar data. The sensing volume of the MRR and radar data is also sketched (Fig. 4). The MRR is a low-power vertical-pointing frequency-modulated continuous wave (FMCW) Doppler radar operating at 24 GHz (Peters et al. 2005). The configuration adopted here has with height range bins of 50-m depth and the measurements reach up to 1500 m above ground. We are using 1-min averaged profiles. From measured Doppler power spectra the raindrop size distribution (RDSD) is estimated in a vertical profile above the radar. First, each spectral velocity bin is assigned a raindrop size using a known size–fall velocity relation. Second, the number of raindrops per size interval is estimated from the spectral reflectivity using the scattering cross section of a single raindrop. At this step Mie scattering of raindrops is assumed. Finally, from the retrieved RDSD the radar reflectivity factor is computed using

$$z_{\text{MRR}} = \frac{\sum N_i(D)D^6}{C^2}$$

(Peters et al. 2005). Here, we use the attenuation-corrected 1-min averages computed by the MRR firmware (Peters et al. 2010). The calibration of an MRR can be monitored using rain gauges (Chen et al. 2015). For absolute calibration, the system is characterized at the manufacturer’s test range using a corner reflector with a known backscatter cross section. A critical factor in the retrieval of the RDSD is the assumption of zero vertical air motion and the absence of turbulent broadening of the Doppler spectrum. Vertical air motion will shift the Doppler spectra, resulting in the wrong assignment of the raindrop size (Hauser and Amayenc 1981). The reflectivity $\eta$ of the observed Doppler spectra is invariant to shifting of the spectra; however, the estimation of the reflectivity factor $z_{\text{MRR}}$ via RDSD estimates is affected by any shifting or broadening of the Doppler spectra. Certainly convective updraft and downdraft will affect the retrieval of $z$ from an MRR. A method to mitigate errors in a convective situation is proposed in Adirosi et al. (2016). In our study this source of error is minimized by confining the statistical analysis to stratiform rain situations where strong updrafts and downdrafts do not play a significant role. Upslope effects caused by a steady flow over or around Mount Hohenpeissenberg, which may cause vertically upward motions, are considered to be negligible, since the mountain rises only about 300 m above the Alpine Foreland. More important, we restrict ourselves to weak wind conditions (horizontal wind speed $< 5$ m s⁻¹).

In this study we consider data from the warm season between April and November 2014.

### 4. Case study

We first investigate an illustrative case study from 26 to 27 July 2014. In total 55.9 mm of precipitation were measured at Hohenpeissenberg observatory in a 13-h period. Substantial rain rates were observed with a total of 25.5 mm within 57 min beginning at 1620 local time (LT), which was noted as the most intense period by the local weather observer. Synoptically, southern Germany was governed by low pressure with a warm and humid air mass. In this air mass, slow-moving thunderstorms
developed that were responsible for the strong rain rates. We consider data from the birdbath scan of the C-band system. The time–height profiles of \( Z \) and Doppler velocity \( u \) for the whole height range of the event are shown in Fig. 5. There, the melting layer is clearly visible from \( u \). Small fall velocities (\(<-2 \text{ m s}^{-1}\)) are observed above the melting layer and large velocities below the melting layer (\(<-5 \text{ m s}^{-1}\)). During this event, the melting layer height increased from about 2000 to 2500 m. The initial convective nature up to about 1800 UTC is indicated by positive vertically upward Doppler velocities (maximum 4 m s\(^{-1}\)) above the melting layer. In Fig. 6 we show a time series of rain rate (based on disdrometer data) and radar reflectivity \( Z_{\text{radar}} \) at the first far-field range bin (radar ff), MRR data \( Z_{\text{MRR}} \) (MRR ff) at 650 m, and the disdrometer data (LNM). Data from 26 to 27 Jul 2014 are shown.

MRR seem to deviate stronger from \( Z \) based on the disdrometer (Fig. 7, middle panel). This is due to the initial convective nature, where it is difficult to match disdrometer and radar data, and larger sampling errors. In addition, the underlying assumptions of zero vertical velocity to determine a DSD from MRR data are violated (Peters et al. 2005). Between 1800 and 2000 LT, the agreement between disdrometer and radar data is significantly better, which relates to the stratiform situation indicated by the absence of updrafts in the measurements (Fig. 6). Differences again become large starting at about 0300 LT, when the precipitation measured by the disdrometer stems only from a shallow cloud not visible to the radars at a range 600 m above the site (Fig. 6).

The reflectivity differences relative to the disdrometer data (\( \Delta Z = Z_{\text{dis}} - Z_{\text{MRR}} \) and \( Z_{\text{dis}} - Z_{\text{radar}} \)) and the corresponding fall velocities (middle panel) and cross-correlation coefficient \( r_{hv} \) (lower panel) are shown in Fig. 7. Again, this refers to the difference between data at 650 m and the surface-based disdrometer data. There is good agreement between MRR and radar fall velocities for large \( r_{hv} \) (larger than about 0.95). Differences are larger for smaller \( r_{hv} \), which corresponds to periods of small \( Z \) and larger temporal variability in precipitation, or they are due to heterogeneous scatters in the sensing volumes.

Because of the configuration, the vertical extent of the MRR is limited to 1600 m above the surface. One motivation of this study was to assess the variability in the vertical profile, in particular between 650 m and the surface. This will be investigated statistically in the next
The radar reflectivities of the MRR and the radar are shown in Fig. 8. In the near-field range of the radar antenna, we find line-shaped structures that should not be interpreted quantitatively. Note that the MRR data have a 1-min resolution compared to the 5-min update of the radar data. The 1-min resolution is sufficiently high to reveal tilted columns of rainbands with increased rain rates. The column of large $Z_{MRR}$ close to the surface around 1610 LT shows possible attenuation with $Z_{MRR}$ decreasing with height. Overall, there is a good match between two sensors when looking at the $Z$ patterns. The same is true for the fall velocities (Fig. 9). Note that MRR fall velocities larger than 9.4 m s$^{-1}$ are not resolved by the MRR (Peters et al. 2005). Inherent to the measurement technique of the MRR, positive upward velocities cannot be derived from the MRR measurements. To conclude, this comparison suggests that the MRR may be used to assess the gap between the surface and the first far-field range bin of the C-band radar when the limitations of the MRR, for example, in convective situations are taken into account.

5. Data analysis

We now use disdrometer measurements to monitor the absolute calibration of a radar. One assumption is that the drop size distribution, and therefore the scattering characteristics between the far-field limit ($\sim 650$ m) and the disdrometer near the surface, does not change significantly. This assumption will now be assessed using the additional data from the MRR, which provides a $Z$ profile down close to the surface. The analysis is split into two sample periods: sample 1 from 6 April to 1 July 2014 and sample 2 from 2 July 2014 to 13 November 2014. On 1 July 2014, the system was recalibrated according to the results of the first phase. In the following we will focus on $Z_h$; similar results are found for $Z_v$, the vertically polarized reflectivity factor. The precipitation measurement of the LNM (1-min average) is related to the radar measurement in the first far-field range bin. Assuming a mean drop fall velocity of 4 m s$^{-1}$, it takes nearly 3 min before the precipitation volume sampled by the radar in the first far-field range bin reaches the surfaces. Here, we use the retrieved MRR fall velocity in order to correct for the time offset $\Delta T_z$. The time offset is $\Delta T_z = z_{ref} \times u_{med}$ with $u_{med} = \text{median}(u_i^{ref})$; $z_{ref}$ is the height of the radar range bin, which is related to the disdrometer measurement; and $u_i$ is the fall velocity at a particular height between the surface and the reference height. Furthermore, we exclude brightband situations using only data with $T > 4^\circ\text{C}$ at 650 m, $\rho > 0.98$, and Doppler velocities $< -2$ m s$^{-1}$. The temperature at radar height is estimated using temperature measurements at 2-m height and assuming a
moist adiabatic temperature gradient. In total two datasets with a precipitation duration of 182 h for sample 1 and 390 h for sample 2 are available for analysis.

We first discuss a scatterplot of $Z$ based on the different instruments before those scatterplots are statistically evaluated. Radar reflectivity $Z$ from the radar and MRR versus the disdrometer data is shown in Fig. 10. The radar–MRR comparison is shown in Fig. 11. The results for the second sample are shown in Figs. 12 and 13.

To compute the biases, we filter the data for fall velocities $\leq -2 \text{ m s}^{-1}$, reflectivities between $15 < Z_{\text{radar}} < 35 \text{ dBZ}$, and $Pr > 0.98$ for which we find the smallest variability in the profile up to 650 m. As such, these cases are considered to define a stratiform situation with the most homogeneous precipitation conditions in the vertical. In general the agreement between disdrometer and MRR data is quite good for both samples. There is a larger spread in the distribution during the second sample, in particular for data $< 15 \text{ dBZ}$ (Fig. 12). The bias, which is defined as the median of the difference between the disdrometer and MRR is 0.1 dB and is essentially constant for sample 1 and sample 2, and the interquartile range (IQR) of the distribution (defined by the first and third quartiles) increases from 4.3 to 5.3 dB (Table 1). This is also indicated by the median absolute deviation (MAD; Wilks 2011), which increases from 1.6 to 2.5 dB for the second sample period (Table 1). MAD is defined as $\text{median} |x_i - q_{0.5}|$, or the median of the absolute difference between the $i$th sample $x_i$ and $q_{0.5}$, which is the median of the sample. The MAD definition has the advantage that it considers the full sample, while outliers do not affect the estimate, as is the case for standard deviation. So, it can be considered as a more robust standard deviation characterizing the spread of the data samples around the mean bias. The reason for this increase of the disdrometer/MRR MAD value in the second sample period is not yet clear. We find a larger number of events with greater rainfall intensity. This in turn may lead to larger attenuation of the MRR and to larger disdrometer sampling errors of large drops, and therefore contribute to a larger MAD value. The mean bias between the C-band radar and the

![Fig. 10](image1.png)  
**Fig. 10.** One-to-one scatterplot between disdrometer d1 ($Z_{\text{dis}}$) and (top) radar data ($Z_{\text{radar}}$) in the first far-field range bin, for the first sample period, and (bottom) MRR ($Z_{\text{MRR}}$).

![Fig. 11](image2.png)  
**Fig. 11.** One-to-one scatterplot between radar ($Z_{\text{radar}}$) and MRR data ($Z_{\text{MRR}}$) in the first far-field range bin for the first sample period.
disdrometer decreases from 1.8 to 0.7 dB, which is due to the adjustment of the calibration based on the first sample phase (a 1.5-dB adjustment of the dBZ0 was applied; dBZ0 is the minimum detectable radar reflectivity at a range of 1 km). The disdrometer/radar MAD values remain essentially constant. The IQR of the distribution decreases from 4.1 to 3.7 dB. For the C-band radar–MRR bias (by definition we use the C-band radar as a reference), the bias decreases from −1.7 to −0.2 dB, the IQR increases from 3.5 to 5.0 dB, and the radar/MRR MAD value increases from 1.6 to 2.5 dB. Apparently, the MRR measurements are noisier in the second sample because this is observed independently for the LNM and C-band radar comparison. These results here are comparable to those found by Tokay et al. (2009).

Comparable MAD values are found when assessing the spatial correlation among the three available disdrometers following Lee et al. (2009); see the appendix. There, we find high spatial correlations, which indicate that the disdrometer d1 (used here) is representative up to a scale of 200 m for this location. The disdrometer/radar MAD values found here represent the inherent spatial variability of the precipitation process. This result justifies the approach of relating a volume measurement to a time-average-based measurement under stationary conditions, by considering only stratiform precipitation events.

From Table 1 we can conclude that the vertical variability of Zhradar and ZMRR is relatively small, because overall the computed biases are rather small. This is also highlighted by the fact that the one-to-one correspondence is rather good (Figs. 10–13). Furthermore, the case study discussed in the previous section shows this for situations with moderate reflectivities, where we find small differences between radar and MRR (Fig. 6; see, e.g., measurements between 1800 and 2000 LT).

The MRR data are now used to quantify the vertical variability in the atmospheric column up to the far-field limit. The MRR-measured ZMRR at 650 m against the column-averaged ZMRR is shown in Fig. 14 (upper panel). There is a good one-to-one relationship, especially for
Z_{MRR} > 10 dBZ. There is a larger scatter below Z_{MRR} < 15 dBZ. From this we can already argue that the good correspondence between the first far-field range bin Z_{MRR} and the column-averaged Z_{MRR} supports the approach to relate the far-field Z_{MRR} and as such Z_{h,radar} to the disdrometer measurements. A measure of the variability in this column is given by the standard deviation of Z_{MRR} in the column at a given Z_{MRR} value in the far field (Fig. 14). Every point in this scatterplot corresponds to the standard deviation of Z_{MRR} for a given MRR profile in relation to the Z_{MRR} at 650 m. We then compute the median standard deviation for 1.5-dB large Z_{MRR} bins as a function of the far-field Z_{MRR} (the black curve). In addition we show the range of the standard deviation given by the first and the third quartile values of the distribution. In a range between about 15 and 35 dBZ, the median of the standard deviation is around 1 dB with a range of ±0.5 dB. So, the variations of Z_{MRR} in the profile typically have a magnitude of 1 dB or smaller. Larger variations are found for smaller Z_{MRR} values (Fig. 14). This may be related to situations where the precipitation process has started to develop. Evaporation effects and a nonequilibrium drop size distribution may cause larger variations in Z_{MRR} (Kumjian and Ryzhkov 2010). The observed small variability of Z_{MRR} with a height in a range between about 10 and 35 dBZ explains why we can relate the measurement at 650-m height with disdrometer measurements near the surface. This implies that the drop size distribution is on average constant with height during these situations. This is consistent with the findings of Peters et al. (2005).

It is well known that disdrometer measurements are sensitive to large horizontal wind speeds (Friedrich et al. 2013). Using data from the collocated sonic anemometer, there are only a very few situations in this dataset where the wind speed exceeds 5 m s⁻¹.

So far we have focused on associated biases in Z between the first far-field range bin of the C-band antenna and the disdrometer. We also compute the mean bias of Z_{MRR} relative to the disdrometer for the MRR at all range gates between the surface and 1600 m. The results for the first and second phases of the measurement campaign are shown in the upper panel in Fig. 15. This has to be compared with the mean bias without performing a time correction as a function of height (lower panel, Fig. 15). Overall, the bias, defined as the median of the difference between LNM and MRR, has in both cases a magnitude of 1 dB. But the spread of the distribution, defined by the first and third quartiles is significantly smaller for the time-corrected error profiles, especially at higher height levels. This illustrates the significance of applying a time correction. In both cases the bias shows a substantial increase in the first range bin. These near-field measurements in the first to three range bins are not reliable and therefore are discarded from the discussion.

To use this monitoring approach in an operational environment, it is important to know how large the sample size needs to be before the bias estimate converges to a stable value. We assume that the bias does not have a trend with time. The Z bias as a function of sample size and for the two sample periods is shown in Fig. 16. The bias for a given sample size is normalized with the overall bias that is computed with the whole data sample. From this a roughly 20-h sample size is needed to be within about 0.5 dB of the final bias results. This sample size means that 240 birdbath scans (one every 5 min) under stratiform precipitation conditions are at least needed for the bias estimate. The 20-h sample size threshold can be found for both the phase 1 and 2 sample periods.

### 6. Monitoring calibration with other sources

In this section we compare the previous analysis to other methods to monitor the radar calibration. We first discuss the result of the calibration scan, which provides a one-point calibration of the receiver twice a day. The computed dBZ0 for the H and V channels (dBZ0h and dBZ0v, respectively) are compared to the system dBZ0 during the period considered (Fig. 17, top panel). Clearly visible is the adjustment of the dBZ0 at the beginning of July. Overall, the dBZ0 in the system and the ones computed from the calibration scan match within ±1 dB. However, there are distinct differences. Roughly in the beginning of May, there is an increase in
dBZ0h by about 1 dB and in dBZ0v by 0.5 dB (the system is losing sensitivity). This relates to a full recharacterization of the system: for all transmit and receive losses, the transmit power was measured and saved to the system, separately for both channels, without adjusting the calibration of the system for this test. A new characterization results into a new radar constant that is responsible for the increase in dBZ0 by about 1 dB at the beginning of May. The dBZ0 values from the one-point calibration using the internal TSG are relatively constant with time. After the recalibration of the system at the beginning of July, we found a bias of 0.7 dB using the disdrometer as a reference (see Table 1). The results from the internal calibration suggests a roughly 1-dB better sensitivity of the radar in H (dBZ0 according to this internal calibration is around $238.8 \text{ dB}$ vs $237.8 \text{ dB}$ as the result of radar–disdrometer analysis; see Table 1). If this were taken over during the calibration, then the consequence would have been an increase in the bias by about 1 dB. With this type of engineering characterization of the radar system, the real bias cannot be quantified accurately enough. Obviously, the uncertainties and biases in determining the components of the radar equations, such as the computation of antenna gain measurement, are too large to be more accurate. The approach for using an external reference (precipitation) has also the advantage that the full transmit and receive path of the signal is monitored so that changes due to the radar hardware can be detected.

Fig. 14. Term Z from the (top) MRR at 650 m vs the mean Z in the column between the surface and 650 m, and (bottom) $Z_{\text{MRR}}$ vs standard deviation of $Z_{\text{MRR}}$ in the column between 650 m and the surface, also based on MRR data. “Filtered” data points are data with fall velocities larger $2 \text{ m s}^{-1}$, based on the C-band Doppler data.

Fig. 15. The Z bias between disdrometer and MRR as a function of (top) height–time corrected and (bottom) no time correction. The spread of the bias is defined by the first and third quartiles of the distribution at a given height. Shown are the results for the two sample periods.
Another method for partially assessing the quality of the calibration is to use solar data. Operationally, the receiver is monitored using radar-measured sun power during operational scanning (typically 30 hits a day). The adapted method, which is based on the work of Holleman et al. (2010), is described in Frech (2013). The solar data are also used to monitor the pointing accuracy. When analyzing the solar power seen by the radar, we can characterize only the receive path, not the transmit path. However, in combination with the above-described approach, it is possible to separate errors with respect to the transmit and receive path. The solar power measurements, extracted from the operational scanning against independent solar power measurements (available three times a day), are shown in Fig. 18. The solar cycle is clearly visible, and the largest differences are found when the sun is active (see, e.g., Holleman et al. 2010). Overall, there are no apparent trends during the two samples (April–November). Most of the time, the difference is smaller than 1 dB. Larger differences can be attributed to situations where the sun shows enhanced activity. For the first sample (6 April–1 July 2014), the mean bias (defined by the difference in solar power–radar-received power) of the H channel is $-0.2$ dB and for sample 2 (2 July–13 November 2014) $-0.3$ dB. These numbers imply that the overall bias of 1.8 dB (Table 1) can be attributed to errors in characterizing the transmit path.

### 7. Conclusions

In this work we have evaluated the monitoring of the radar calibration using disdrometer measurements in close vicinity of a weather radar, where we relate the
disdrometer data to radar data from a birdbath scan, which is available every 5 min. One key assumption is that radar measurements in the first far field of the radar antenna (about 650 m above the radar) can be related disdrometer data taken close to the surface. An MRR is used to close the gap between the first radar range bin of the radar and the disdrometer close to the surface. We show that the variability of $Z$ between the far-field limit of the radar and the surface is small if we avoid brightband effects and confine the analysis to stratiform situations that are identified with mean drop fall velocities $< -2 \text{ m s}^{-1}$, reflectivities between $15 < \text{Z}_{\text{radar}} < 35 \text{ dBZ}$, and $\rho_{hv} > 0.98$. The bias between radar and disdrometer data is initially 1.8 dB. There is no bias between the MRR data at 650 m above the surface and the disdrometer data. Based on the results of the disdrometer–radar comparison, the calibration of the radar is adjusted. This reduces the bias to almost 0 dB. The need to adjust the calibration is not reflected by the internal one-point calibration even though a very careful characterization of the system (losses, transmit power) was carried out. This implies that the classic calibration approach has its limitations and that external data sources are needed to quantify the calibration bias. The representativeness of the disdrometer is assessed using two additional disdrometers. We find high spatial correlation coefficients ($> 0.9$) and comparable MAD values, which supports the use of a disdrometer for monitoring the calibration of a radar with this setup employing the birdbath scan (available every 5 min). Solar monitoring is used to monitor and characterize the receiver. The mean bias is for both samples $-0.2$ and $-0.3$ dB, where the sun is considered as a reference. Those small numbers indicate that the largest contribution to the observed bias presumably can be found in the transmit path.

We have shown a case study of an intense precipitation event that was of convective nature initially. Qualitatively, the agreement between MRR and radar data is remarkably well considering the spatial and temporal structures of the radar reflectivity and the HM fall velocities. Quantitatively, the agreement of the reflectivities is very well for values between about 20 and 30 dBZ. Larger differences are found when the meteorological situation becomes heterogenous in time and space. So depending on the situation, MRR measurement assumptions may be violated, which may lead to biases in the estimated moments. Furthermore, inherent sampling issues become larger when relating a point measurement to a volume measurement.

The combination of disdrometer and birdbath radar data is set up in a way that it is operated easily in an operational environment. So far, six radar sites are equipped with disdrometers. There are plans to install a disdrometer at all sites. The installation will also include a sonic anemometer for better quality control of disdrometer data. This monitoring approach is one component of the overall monitoring of the data quality, which includes conceptually more sources [the multisource radar calibration (MuSoRCa) concept]. It includes the evaluation of solar spikes in radar data, radar–radar comparisons, and a close monitoring of the radar system state. It is also planned to include the self-consistency approach into this framework in the future. A decision tree still needs to be defined based on which calibration is eventually adjusted automatically. In addition, the proposed approach needs to be extended to winter weather situations so it can serve as a year-round monitoring tool.

Acknowledgments. The discussions with Jörg Seltmann are greatly acknowledged. The comments of the anonymous reviewers helped to improve this paper.
APPENDIX

Disdrometer Analysis

A point measurement (disdrometer) and the corresponding time averaging is related to a volume measurement. It is known that the variance between a point and a volume measurement can be substantial (Lee et al. 2009; Thurai et al. 2012). Considering the spatial separation here (200 m), this error is expected to be small. Since the instruments are located on a mountain top, horizontal homogeneous conditions may be questioned, in part because of the induced flow over a mountaintop. The three available disdrometers (see Fig. 3) can be analyzed in terms of systematic differences in order to assess the representativeness of disdrometer d1, which is used for the calibration analysis.

For a given time \(i\), we compute the difference \(\Delta Z_i = Z_{\text{dis1},i} - Z_{\text{dis2},i}\), where d1 and d2 denote the respective disdrometers. The scatterplots are shown in Figs. A1 and A2. The median, first and third quartiles, MAD, and the correlation of \(Z\) and the rain rate time series are summarized in Table A1. This is also done for the measured rain rate. Furthermore, we compute the "noncentered" correlation of spatially separated variables, which avoids subtraction of the mean (Lee et al. 2009) and provides a measure of space independence for a nonstationary process:

\[
\rho_{d_i,d_j} = \frac{\sum d_i(x_1,y_1,t)d_j(x_2,y_2,t)}{\sqrt{\sum d_i(x_1,y_1,t)^2 \sum d_j(x_2,y_2,t)^2}}^{0.5}, \quad (A1)
\]

where \(d_i, d_j\) denote the disdrometer measurements (\(Z\) or the rain rate) at time \(t\) of a pair of disdrometers at a location \((x_1, y_1)(x_2, y_2)\), respectively.

The overall \(Z_{\text{dis}}\) correlation between the two instruments is about 0.8; the median difference between the two sensors is 0.6 dB for \(Z_{\text{dis}} > 10\) dB and 0.3 dB if we consider only data \(20 \leq Z_{\text{dis}} \leq 30\) dB. For the same \(Z_{\text{dis}}\) interval, the bias between LNM and Parsivel2 is \(-0.7\) dB, meaning the Parsivel2 computes on average larger \(Z_{\text{dis}}\) values. But still the agreement between the sensors is remarkably good. The MAD is smallest for the two sensors d1 and d2, which are separated by about 20 m. The MAD increases from 1.5 to about 1.8 dB when relating disdrometer d3 to disdrometers d1 and d2. This increase is attributed to the spatial variability of

FIG. A1. One-to-one scatter of the \(Z_{\text{dis}}\) plot between disdrometers d1 and d3, which are separated by approximately 200 m (July–November 2014).

FIG. A2. One-to-one scatterplot of the \(Z_{\text{dis}}\) plot between disdrometer d1 and d2, separated by approximately 20 m (July–November 2014).

TABLE A1. Consistency of three disdrometers: d1 (Thies), d2 (Parsivel2), and d3 (Thies, 200 m away from d1); see Fig. 3. Differences of \(Z\) (dB) are statistically analyzed. The first and third quartiles, and the median differences of \(Z\) and MAD are shown. Also given are the correlations \(\rho_Z\) and \(\rho_{\text{RR}}\) of the disdrometer measurements for \(Z\), and the rain rate. Data are from July to November 2014.

<table>
<thead>
<tr>
<th></th>
<th>d1–d3</th>
<th>d1–d2</th>
<th>d3–d2</th>
</tr>
</thead>
<tbody>
<tr>
<td>First quartile (dB)</td>
<td>1.2</td>
<td>2.2</td>
<td>3.17</td>
</tr>
<tr>
<td>Median (dB)</td>
<td>0.6</td>
<td>0.7</td>
<td>1.2</td>
</tr>
<tr>
<td>Third quartile (dB)</td>
<td>2.4</td>
<td>0.9</td>
<td>0.8</td>
</tr>
<tr>
<td>(\rho_{Z,d_i,d_j})</td>
<td>0.92</td>
<td>0.82</td>
<td>0.82</td>
</tr>
<tr>
<td>(\rho_{\text{RR},d_i,d_j})</td>
<td>0.96</td>
<td>0.88</td>
<td>0.88</td>
</tr>
<tr>
<td>MAD</td>
<td>1.8</td>
<td>1.5</td>
<td>1.9</td>
</tr>
</tbody>
</table>
precipitation. The high correlation among the three sensors (nearly 0.9 or better for the rain rate, and somewhat less for Z, which is expected because it represents a higher-order moment; see Lee et al. 2009) indicates remarkably small spatial variability of the disdrometer measurements for our setup. The computed MAD has the same magnitude as we find in the radar–disdrometer results. This reflects the inherent spatial variability of the precipitation process. But the high spatial correlation found in the disdrometer measurements suggests a reliable radar bias estimate can be derived from this setup. A dedicated calibration experiment is needed (see, e.g., de Moraes Frasson et al. 2011) to assess the observed differences, especially between the Thies and Parsivel disdrometers.

REFERENCES


