15 Sep 2019

Inter-comparison Of Reflectivity Measurements Between GPM DPR And NEXRAD Radars

Munsung Keem

Bong Chul Seo
Missouri University of Science and Technology, bongchul.seo@mst.edu

Witold F. Krajewski

K. Robert Morris

Recommended Citation

M. Keem et al., "Inter-comparison Of Reflectivity Measurements Between GPM DPR And NEXRAD Radars," *Atmospheric Research*, vol. 226, pp. 49 - 65, Elsevier, Sep 2019.
The definitive version is available at https://doi.org/10.1016/j.atmosres.2019.04.010

Atmospheric Research 226 (2019) 49–65

Contents lists available at ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmosres

Inter-comparison of reflectivity measurements between GPM DPR and NEXRAD radars

Munsung Keem[a,*], Bong-Chul Seo[a], Witold F. Krajewski[a], K. Robert Morris[b]

[a] IIHR—Hydroscience & Engineering, The University of Iowa, Iowa City, IA, USA
[b] Science Applications International Corporation (SAIC), Reston, VA, USA

ARTICLE INFO

Keywords:
GPM DPR
Ground radar calibration
Partial beam blockage

ABSTRACT

This study demonstrates the potential use of the NASA's Global Precipitation Measurement (GPM) Dual-frequency Precipitation Radar (DPR) to examine ground radar (GR) miscalibration and other uncertainty sources (e.g., partial beam blockage). We acquired the GPM Ground Validation System Validation Network reflectivity matchups between the DPR and three GRs (two in Iowa and one in South Dakota) for 2014–2017. We then refined the matching parameters (e.g., time separation) to reduce uncertainty in the matchup samples by analyzing the sensitivity of the matchup statistical properties to changes in these parameters. To reconcile the same observables (i.e., reflectivity) with different observational properties among the space- and ground-based radars, we developed a statistically integrated framework using inter-comparisons of them all with a Monte Carlo simulation. This method verifies the absolute calibration bias estimated from the refined DPR–GR matchups using relative calibration biases between GRs. We found that taking samples with a narrow temporal gap, estimated by actual measurement time of the DPR and GRs, can significantly reduce sample variability. Through inter-comparisons among the DPR and GRs, we observed that reflectivity differences among GRs in a similar environment (e.g., climatology and geography) are likely to be affected primarily by the calibration mismatch. In this case, the inter-comparison results demonstrated good agreement, and we inferred that the differences can be mitigated by calibration bias correction against the DPR. On the other hand, when the disagreement level of the inter-comparison results is significant, the authors found that other factors, such as partial beam blockage even in relatively plain regions, are more dominant than the calibration bias. In fact, the partial beam blockage effects can manifest themselves as a seasonal pattern in the GR inter-comparison results.

1. Introduction

In current ground-based precipitation estimation systems, weather radar (hereafter Ground Radar, GR) observations play a central role in determining the spatial and temporal structures of precipitation systems. Their advantages are highlighted in real-time weather and flood prediction, where rapid updates of precipitation information over broad spatial domains are required. However, our incomplete understanding of the weather system under various environmental conditions serves as the main obstacle to eliminating a variety of uncertainty sources in Quantitative Precipitation Estimation (QPE) (e.g., Villarini and Krajewski, 2010).

The types and magnitude of the uncertainties involved in different GRs can vary (e.g., Bringi et al., 2011; Thurai et al., 2017). The differences often manifest themselves on large-scale rainfall accumulation maps derived from multiple radar observations as discontinuous features at the equidistance zones among radars. For example, Fig. 1 illustrates a rainfall accumulation map for the State of Iowa from 1 July to 31 August 2017, derived from the Iowa Flood Center (IFC) real-time radar QPE (Krajewski et al., 2017). In Fig. 1, we can recognize clear border lines at the overlapping areas among the involved radars, particularly centered on the Sioux Falls radar (KFSD) in South Dakota. We have also observed similar patterns in other radar composite products (e.g., the Multi-Radar Multi-Sensor product; Zhang et al., 2016). One plausible explanation for this issue might be the negative calibration bias of the KFSD radar (i.e., underestimation) because we can observe the same precipitation at equidistant zones; all areas covered by the KFSD radar show this underestimation tendency, and there are no significant topographic features (i.e., mountain effects) in Iowa. Recently, Zhong et al. (2017) and Warren et al. (2018) reported the same issues existing in the radar networks in China and Australia, respectively. Many studies have identified radar miscalibration as one of the most

* Corresponding author at: IIHR—Hydroscience & Engineering, The University of Iowa, 100 C. Maxwell Stanley Hydraulics Laboratory, Iowa City, IA 52242, USA.
E-mail address: munsung-keem@uiowa.edu (M. Keem).

https://doi.org/10.1016/j.atmosres.2019.04.010
Received 9 December 2018; Received in revised form 2 April 2019; Accepted 9 April 2019
Available online 11 April 2019

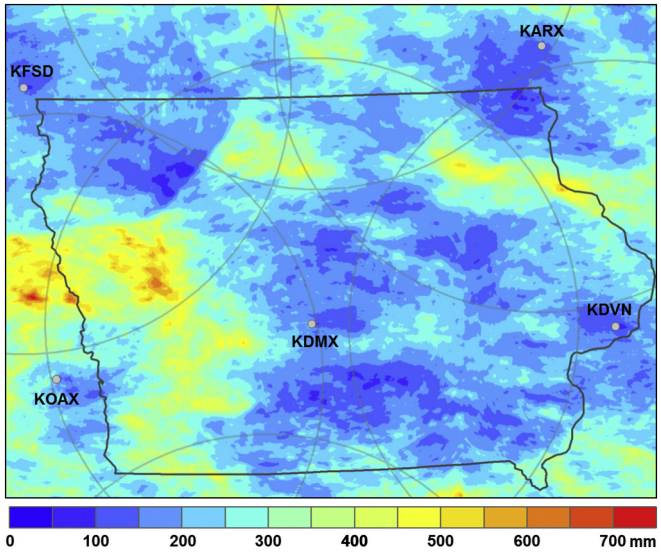

Fig. 1. A rainfall accumulation map derived from the IFC radar-only products over the state of Iowa from 1 July to 31 August in 2017.

significant practical issues affecting radar uncertainties (Smith et al., 1996; Brandes et al., 1999; Wang and Wolff, 2009; Seo et al., 2014; Zhong et al., 2017). Small discrepancies in calibration may lead to large uncertainties in QPE, as well as hydrometeorological applications, because of their nonlinear propagation properties (e.g., Seo et al., 2013).

The literature documents many methods to calibrate GRs using a variety of targets (e.g., a metal sphere and corner reflectors) and independent observation systems (e.g., rain gauges and disdrometers) (e.g., Atlas, 2002; Frech et al., 2017). However, these methods may not be suitable to detect abrupt changes (in real-time) and/or gradual degradation of the (calibration) offset caused by decreased performance of the radar hardware system over time (Anagnostou et al., 2001). Alternatively, we can use direct comparison of reflectivity data among adjacent GRs for the same targets (Smith et al., 1996; Seo et al., 2013, 2014). The main advantage of this method is that it can benefit from the high temporal resolution of GR observations, and the comparison results can be immediately applied to bias adjustment. In this case, other uncertainty sources such as anomalous beam propagation and range effects might be compounded together with the calibration error; selecting appropriate samples should be done with great care to isolate the calibration bias. Furthermore, the lack of detailed information on how much adjustment and when the GR calibration was performed makes it challenging to verify the accuracy of this method (Seo et al., 2014).

Space-borne radars—in particular, the Ku-band precipitation radar (PR) onboard the Tropical Rainfall Measuring Mission (TRMM) satellite and the dual-frequency (Ku- and Ka-band) precipitation radar (DPR) onboard the Global Precipitation Measurement (GPM) mission *Core Observatory*—have also received attention as a tool for monitoring GR's performance (Anagnostou et al., 2001; Liao et al., 2001; Wang and Wolff, 2009; Schwaller and Robert Morris, 2011; Gabella et al., 2013; Hou et al., 2014; Speirs et al., 2017; Zhong et al., 2017; Warren et al., 2018). The PR/DPR can serve as a reference to estimate the calibration bias of GRs (Schwaller and Robert Morris, 2011) because it not only produces the same variable (i.e. reflectivity) with stable calibration level (± 1 dB), but it can also be free from a substantial portion of the uncertainties common in GR products, such as ground clutter and beam blockage effects. Previous studies asserted that the accuracy of GR's precipitation estimates could be improved by correcting their calibration biases against space-borne radars. However, the low temporal frequency of space-borne radar measurements leads to their limited usage for operational applications.

The main purpose of this study was to examine reflectivity disagreement among adjacent GRs using relatively consistent (in space) measurements from the GPM. We speculate that the issue shown in Fig. 1 is mainly caused by differences in GR calibration offsets, which raises the central question of this study: are the reflectivity differences observed among GRs caused by miscalibration? To address this question, we compared the ground-based and space-based approaches to evaluate and diagnose GR's calibration bias, while taking advantage of supplementary aspects of the DPR and GRs: (1) high temporal resolution of GRs; and (2) spatial consistency and stable calibration of the DPR.

To reconcile the same observables (i.e., reflectivity) with differing observational properties among the space- and ground-based radars, we developed a statistically integrated framework with a Monte Carlo simulation. The proposed method is validated through the inter-comparisons among the GPM DPR and GR products using the GPM Ground Validation System (GVS) Validation Network (VN) dataset and a reflectivity comparison method for adjacent GRs (developed in this study). We expect the use of reflectivity measurements in the comparison to prevent the uncertainty associated with the variability of Z–R relationship. In addition, this approach allowed us to avoid the assumption of zero mean errors for DPR measurements, which may not be true but has been used widely by many other studies (e.g., Anagnostou et al., 2001; Wang and Wolff, 2009; Schwaller and Robert Morris, 2011; Kim et al., 2014; Warren et al., 2018). Instead, we assumed that only the DPR's error distribution is stationary. We also noted that the approach of Warren et al. (2018) is similar to the one proposed in this study; both evaluated space-borne radar's capability to correct GR's calibration bias by looking at reductions in the reflectivity differences among GRs. However, this study incorporates the reflectivity differences into the examination process on GR miscalibration and investigates other uncertainty sources, if any, that would affect those differences.

2. Data sources

In this study, the GPM GVS VN datasets and individual radar reflectivity products of the IFC are collected for three “NEXRAD” WSR-88D (Weather Surveillance Radar-1988 Doppler) radars in Des Moines (KDMX) and Davenport (KDVN), both in Iowa and Sioux Falls, South Dakota (KFSD), for the period of March 2014 through September 2017 (Fig. 2). We provide more details on each dataset in the following subsections.

2.1. The GPM GVS VN dataset

Since the launch of the GPM mission *Core Observatory* satellite in February 2014 (Speirs et al., 2017), the National Aeronautics and Space Administration (NASA) has been operating the GPM GVS VN to improve the accuracy of the satellite products through comparisons with GR observations. The GPM GVS VN provides several types of geometrically matched datasets (matchups) between GPM satellite and GR measurements using 75 WSR-88Ds as well as various international partners' radars (National Aeronautics and Space Administration, 2015). All input and output datasets used in the VN are accessible through a VN data server (ftp://hector.gsfc.nasa.gov/gpm-validation/data/). Among the datasets, we used the geometry-matched products between DPR and GR (referred to as the DPR–GR matchups) for the comparisons.

The DPR–GR matchups are directly generated from the GPM DPR standard Level 2 products (2ADPR) and coincident GR products (National Aeronautics and Space Administration, 2015), which are averaged over common volumes defined by a matching algorithm (see Section 3.1). The DPR standard products consist of three datasets that depend on the swath types: (1) normal scan (NS); (2) matched scan (MS); and (3) high-resolution scan (HS). Consequently, the DPR–GR

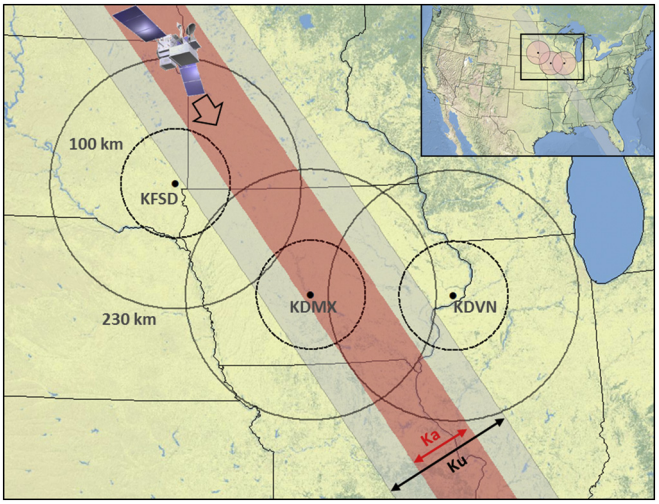

Fig. 2. Study area map. The black points indicate the locations of the selected three NEXRAD radar sites: KDMX in Des Moines, IA; KDVN in Davenport, IA; and KFSD in Sioux Falls, SD. The dotted and solid line circles demarcate 100 km and 230 km range rings from each radar site, respectively. The gray- and pink-shaded areas indicate an example swath of the GPM DPR normal scan (NS) and high resolution scan (HS) over the study area. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

matchups offer the same categories. To ensure adequate sample size, we used the NS matchups, which have the widest spatial coverage. We note that measured reflectivity values of the NS matchups were derived from Ku-band radar reflectivity with attenuation correction. We refer to the attenuation-corrected DPR reflectivity as "DPR reflectivity." A NASA staff specialist performed quality control on the GR reflectivity data in the matchups. The DPR–GR products were generated only when sufficient rainfall matchup samples (≥100 Ku-DPR rays tagged as "rain certain") were detected within 100 km from a GR.

A recent upgrade of the DPR Level 1 algorithm (Version 5) changed the Level 2 products (e.g., matchups). Among all the changes, adjustments in the offset parameters related to hardware performance (such as transmitting power, receiver gain, beam width, and the pulse width) affected the results of this study. We found that the DPR reflectivity values increased (1.17 dB), and 7.6% of precipitation classification results changed for the period of May 2014 through April 2017 within the KDMX domain. The increase in the DPR reflectivity value resulted in precipitation type changes because one of the decision factors for convective rain is a reflectivity threshold of 39 dBZ. Because further investigations of the version change are beyond the scope of this study, we selected the latest version (version 5) matchups for our analyses.

2.2. IFC individual radar products

To compare the reflectivity data between two adjacent GRs, we used the IFC's individual radar-rainfall products. The IFC generates time-synchronized products every five minutes for seven NEXRAD radars covering Iowa; these serve as the main input for the IFC's real-time flood forecasting system (Krajewski et al., 2017). Various studies have used and validated these products (Ayalew et al., 2014; Cunha et al., 2015; Quintero et al., 2016). The basic algorithms used to generate the products are documented in Seo et al. (2011) and have undergone several improvements (e.g., Seo et al., 2015). The product's features include: (1) vertical averaging based on a non-parametric weighting scheme to build the Constant Altitude Plan Position Indicator (CAPPI) at 1.5 km altitude above mean sea level (AMSL); (2) temporal synchronization (every 5-min nominal time) among involved radars that transfer observations at different times; and (3) spatial synchronization onto a common geographic grid with the resolution of 15 arc-second in latitude and longitude (approximately 400–500 m in Iowa). These features enabled us to readily investigate the reflectivity differences between adjacent GRs.

3. Methodology

In this section, we define several key terms frequently used in our analysis: (1) Calibration Bias (CB) — the average difference between a GR and the DPR (reference); (2) Relative Calibration Bias (RCB) — the average difference in CB between two GRs; and (3) Relative Bias (RB) — the average difference for the same meteorological targets between two GRs. In this study, we defined the null hypothesis (H_0): the RCB between two GRs is the same as the RB between the involved GRs if the CB is a dominant factor of RB in Eq. (1); and the alternative hypothesis (H_1): otherwise, other factors would be the main sources of RB.

$$H_0: RCB_{GR1-GR2} = RB_{GR1-GR2} \text{ vs. } H_1: RCB_{GR1-GR2} \neq RB_{GR1-GR2} \quad (1)$$

In the following sub-sections, we describe the estimation methods of each defined element and formulate them to test the null hypothesis.

3.1. Calibration bias (CB) and relative calibration bias (RCB)

The stable calibration of the GPM DPR makes it possible to estimate the CB of multiple GRs with the same reference. In this study, we defined the GR's CB (dB) as the spatially averaged reflectivity difference between the GPM DPR and coincident GR observations.

$$CB_{GR-DPR} = \frac{1}{n}\sum_{i=1}^{n}[Z_{GR}(i) - Z_{DPR}(i)] \quad (2)$$

where GR's and DPR's reflectivity measurements (dBZ) of a DPR–GR matched volume i are expressed by $Z_{GR}(i)$ and $Z_{DPR}(i)$, respectively. The number of matched volumes is denoted by n.

Aside from the radar frequency differences, the main challenge to estimating the CB is beam geometry (e.g., sampling volume resolution) difference and observation time mismatch between the DPR and GRs. To resolve this issue, NASA adopted a geometric volume matching scheme based on the methods developed by Bolen and Chandrasekar (2003) and provided matchup samples. The matched volume is defined as the intersection between a DPR ray and individual GR elevation sweeps with an assumption of standard radar beam propagation. Fig. 3 demonstrates the matching scheme. Spatial resolutions of the DPR are 250 m in the vertical and about 5.2 km (at nadir) in the horizontal. The resolutions of GRs are 250 m in range and 0.95° in azimuth and elevation, oversampled to 0.5° in azimuth. The matching volume size is determined by the DPR field of view coverage (horizontal size) and the vertical extent of GR range bins within the coverage (vertical size). For one matching volume, two mean reflectivity values are computed as the horizontal average of the GR products for the GR and the vertical average of the DPR products for the DPR. The DPR–GR matchup dataset contains information on multiple factors that determine the agreement of different sampling volumes (e.g., partial beam filling ratio, precipitation type, and time separation; see Table 1). We refined these factors and performed a sensitivity analysis to reduce uncertainties in the CB estimation (Section 4.1.1).

We presented the CB estimate at each time with 95% confidence interval. Based on the central limit theorem, the sampling distribution of the CB estimate, which is unknown, follows roughly a normal distribution when the sample size is large enough (Mikosch and Kallenberg, 1998). The confidence interval can simply be estimated by using Student's t-test statistic (Anagnostou et al., 2001; Wang and Wolff, 2009). From the CB estimates of two GRs, we can quantify the RCB:

$$RCB_{GR1-GR2} = CB_{GR1-DPR} - CB_{GR2-DPR} \quad (3)$$

where $RCB_{GR1-GR2}$ (dB) is the CB difference between GR1 and GR2.

Although we cannot directly estimate the RCB at an instantaneous

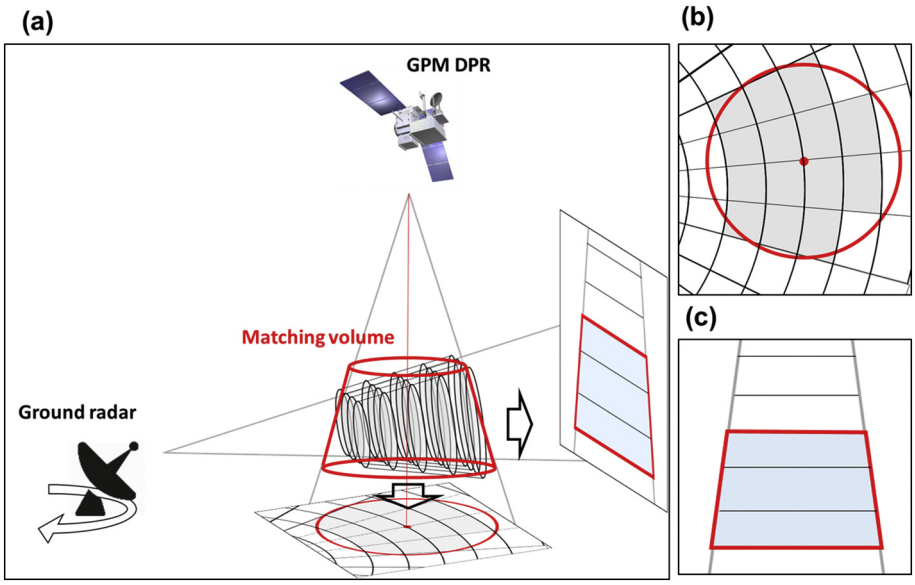

Fig. 3. Schematic representation of the matching volume computation between the GPM DPR and a GR: (a) a quasi-3D schematic of the intersection between a single DPR ray and a single GR sweep, (b) horizontal and (c) vertical cross sections of the intersection.

time (there is no coincident DPR measurement for two GRs at the same time within the 100-km range limits), we use a temporal averaging method to obtain the RCB. We provide further details in Section 3.3.

3.2. Relative bias (RB)

We can evaluate the calibration bias by the direct comparison of reflectivity measurements at overlapping zones between two GRs. Occasionally, this approach may not yield results that are contributed by radar miscalibration, only because other uncertainty sources might be blended in the estimated differences. Therefore, we defined the direct comparison results of GRs' reflectivity products as the "relative" bias, RB. We wanted to determine whether the observed RB is mainly contributed by the relative difference in calibration (i.e., RCB) between GR sites. For this purpose, we used a volume matching strategy between GRs to minimize contributions from other uncertainties. We estimated the RB using the IFC individual radar products (with the minimum reflectivity threshold of 10 dBZ), comparing the averaged reflectivity values included in a specific large volume in the overlapping zones of two GRs (referred to as the Equidistance Sampling Volume, ESV). We note that the radar reflectivity comparison tool (RRCT: http://rrct.nwc.ou.edu/) uses a similar method to provide the relative bias information among adjacent radars over the national network. For estimation of the RB, the main difference is that we compare reflectivity values projected onto common spatiotemporal coordinates. This approach minimizes the error arising from the temporal mismatch of sampling volumes between different radars. We selected the KDMX radar as a reference GR for the comparisons of GRs because it has overlapping zones with both KDVN and KFSD.

We set up the ESV by matching its center line to the equidistance

Table 1
Descriptions of main variables extracted from the DPR–GR geometry matched data (Version 5 2ADPR, NS swath) of GPM GVS VN utilized in this study. More detailed information can be found from GPM GVS VN Data Product User's Guide (National Aeronautics and Space Administration, 2015). All variables are assigned to each matchup.

Variable name	Description	Related parameter
ZFactorCorrected	Attenuation corrected DPR reflectivity (dBZ)	–
GR_Z	Quality controlled GR reflectivity (dBZ)	–
TypePrecip	DPR precipitation type (stratiform/convective/other)	Precipitation type
n_gr_z_rejected	The number of GR bins below 15 dBZ	Partial beam filling ratio (PBFR)
n_gr_expected	The total number of GR bins	
n_dpr_corr_z_rejected	The number of DPR bins below 15 dBZ	
n_dpr_expected	The total number of DPR bins	
site_elev	Ground radar site elevation above mean sea level (km)	Proximity to the bright band
BBheight	DPR bright band height above mean sea level (m)	
topHeight	Top height of a matched volume above ground level (km)	
bottomHeight	Bottom height of a matched volume above ground level (km)	
site_lat	Latitude of ground radar site	Temporal gaps of measurements between DPR and GR
site_lon	Longitude of ground radar site	
latitude	Latitude of DPR surface bin	
longitude	Longitude of DPR surface bin	
scanNum	DPR scan number (zero-based)	
rayNum	DPR ray number (zero-based)	
elevationAngle	Elevation angle of GR sweep (degree)	
timeNearestApproach	Nearest approach time of DPR to GR site (seconds since 1/1/1970 00:00:00)	
timeSweepStart	Starting time of GR sweep for each elevation (seconds since 1/1/1970 00:00:00)	
GR_file	GR Level2 filename	
DPR_2ADPR_file	2ADPR filename	
GR_blockage	Ground radar blockage fraction	–

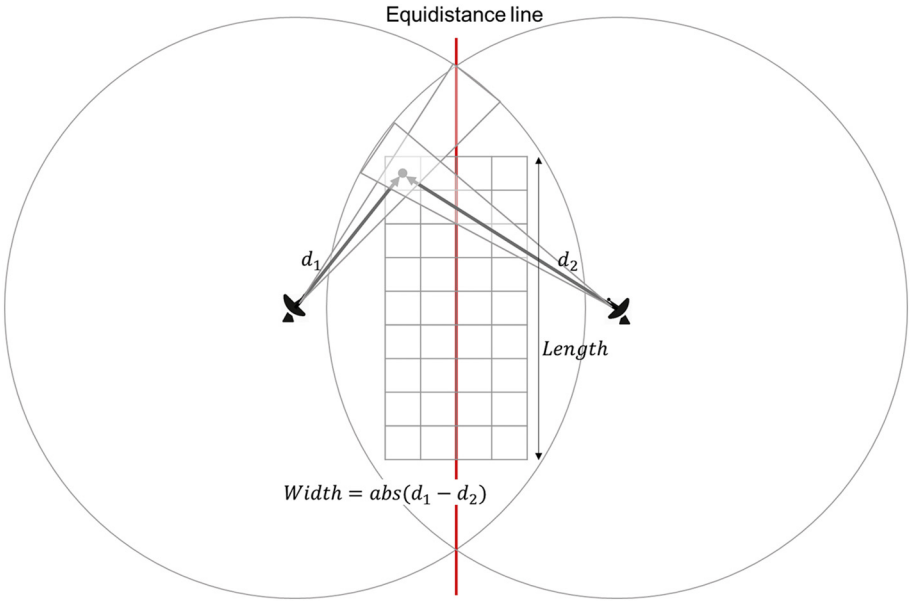

Fig. 4. Schematic representation that demonstrates the construction of the Equidistance Sampling Volume (ESV). Please note that due to the definition of the width parameter, abs($d_1 - d_2$), the actual width of the ESV is not constant but varies along the ESV's length.

line (Fig. 4). We used width and length parameters to adjust the shape of the ESV depending on the distance between GRs. The width parameter is defined as the absolute difference of two distances between each grid center in the ESV and two GRs (d1 for GR1 and d2 for GR2). In fact, it is not equivalent to the actual width of the ESV which varies along the length of the ESV. This parameter plays a role in maintaining range effects at a similar level from both radars so as to mitigate the effects on the reflectivity comparisons. We fixed the width parameter at 5 km for any combinations of GRs. The length parameter was estimated so that the number of grid points within the ESVs for different GR combinations are balanced. Using the ESV allowed us to accelerate computation of reflectivity differences and minimize effects from spatial distortion caused by map projections onto Cartesian coordinates. In addition, the similar number of grids within the ESVs leads to a fair statistical comparison between different GR combinations. Based on the fixed grid system, we decided on length parameters of 171 km for KDMX-KFSD and 144 km for KDMX-KDVN, resulting in 6657 and 7081 grid points in the ESVs, respectively.

Our assumption of the standard beam propagation occasionally fails because of non-homogenous atmospheric conditions along different propagation paths (from both radar sites). This results in the spatial mismatch of rain echoes and the effect of this mismatch on GR intercomparisons might be significant when only small area within the ESV is covered by precipitation. To mitigate this error, we eliminated matchup samples: (1) that have difference > 8 dB (used in the RRCT analysis as an adaptable parameter); and (2) of which the sample size at each observation time is < 10% of the total number of grids in the ESV. Then, instantaneous RB estimates were averaged over a precipitation event. Precipitation events were identified with an arbitrary inter-event time threshold of 12-h. Any consecutive events separated by < 12-h are combined in to a single event. Because radar maintenance (e.g., calibration) is likely not performed during a precipitation event, this event-based RB estimation is reasonable while mitigating sample noise (hereinafter, the event-based RB is denoted by RB). We computed the RB as

$$RB_{GR1-GR2} = \frac{1}{N}\sum_{t=1}^{N} 10\log_{10}\left\{\sum_{i\in V} 10^{Z_{GR1}(i,t)/10} / \sum_{i\in V} 10^{Z_{GR2}(i,t)/10}\right\} \quad (4)$$

where $RB_{GR1-GR2}$ is the event-based RB estimate (dB) between GR1 and GR2. The number of observation times within a precipitation event is denoted by N. The ESV and a pixel in the ESV are expressed by V and i, and $Z_{GR}(i,t)$ is GR's reflectivity (dBZ) at the i^{th} pixel at time t. The form, $\sum_{i\in V} 10^{Z_{GR}(i,t)/10}$, can be interpreted as GR's total reflectivity in linear units ($mm^6\,m^{-3}$) within the ESV. In this study, GR1 is KDMX and GR2 is either KDVN or KFSD.

3.3. Comparison of RCB and RB

The ideal condition to test our null hypothesis using the estimated CB and RB is to have three independent measurements (for the same target) of DPR, KDMX, and another GR. However, the DPR–GR matchups are generated for a significant rainfall event within a 100 km radius from a corresponding GR such that no coincident zones between adjacent GRs exist in these data in our RB estimation domain. Alternatively, we tested the assumption that the RB is caused by the RCB, by averaging Eqs. (3) and (4) over time. Under the null hypothesis, we obtained Eq. (5):

$$E[RB_{GR1-GR2}] = E[RCB_{GR1-GR2}] = E[CB_{GR1-DPR}] - E[CB_{GR2-DPR}] \quad (5)$$

Eq. (5) provides an integrated framework that enables the joint evaluation of three independent bias estimates generated by different platforms at different times. In this equation, it is implicitly assumed that an event-based RB is statistically compatible with a CB estimate. The much narrower comparison area of GR–GR (i.e., the ESV) than that of DPR–GR could be compensated for to some extent by temporally averaging instantaneous RB estimates. This conceptual conversion of time into space may justify the assumption. Before we estimated the expectations in Eq. (5), we needed to verify the stationarity of RB or RCB time series. We tested the stationarity using a nonparametric multiple change point analysis called *E*-Divisive (Matteson and James, 2014). This method is based on a permutation test and provides the number and location of the distributional change points simultaneously without any assumptions of the samples and test statistics distributions. Because we have more frequent samples of RB than of RCB, we performed this analysis on the RB estimates. We can use the results not only to test the stationarity assumption of the RB, but also to assess the DPR's ability to capture the GRs' calibration consistency over time. We used the RCB and RB estimates for the time duration when the stationarity assumption is verified for the null hypothesis tests.

Given a time period during which the stationarity assumption holds, testing the hypothesis presents another difficulty: the different

uncertainty levels in the mean RCB and RB estimates. These are caused by not only unbalanced sample sizes of the three components, $RB_{GR1-GR2}$, $CB_{GR1-DPR}$, and $CB_{GR2-DPR}$, in Eq. (5) but also inherent differences in sample variability due to the different comparison methodologies. We can address this issue with a Monte Carlo simulation as a way of inferring possible RCB behaviors based on the RB estimates. We first assumed the amount of information included in unsynchronized $n_{GR1-DPR}$ (for GR1) and $n_{GR2-DPR}$ (for GR2) CB estimates for the mean RCB estimation is greater than $n_{GR1-DPR}$ and less than $n_{GR2-DPR}$ when $n_{GR1-DPR} < n_{GR2-DPR}$. To minimize the possibility of failure in rejecting a false null hypothesis (type II error), we defined the equivalent sample size of the RB estimates corresponding to the mean RCB as $n_{eq} = \max[n_{GR1-DPR}, n_{DR2-DPR}]$. Then, we took random samples with the sample size of n_{eq} from the RB estimates and averaged them to simulate a sample mean ($\hat{\mu}_i$) of the RB values. We performed this simulation 1000 times and tested the statistical significance of the null hypothesis based on the simulation envelope defined by the minimum and maximum values of $\hat{\mu}_i$. Please note that we decided to use the simulation envelope concept along with the Monte Carlo simulation instead of using the traditional *p*-value approach to test the null hypothesis. Our simulation experiment showed that even when the mean RCB value is estimated from a known population, the p-value can be very small (e.g., < 0.005), due to the unsynchronized and small number of CB estimates. Although the estimated RCB approaches the true value as the sample size increases, this was not the case in our study. Under this situation, the simulation envelope could be more reasonable than the p-value based test.

4. Results and discussion

4.1. CB (DPR vs. GR)

4.1.1. Sensitivity analysis

The DPR–GR matchups from the VN are insufficient to directly quantify the CB of GRs because of probable uncertainties contained in the matchup samples. The main reasons are attenuation of the DPR signal, differences in detectable ranges in reflectivity, high variability of precipitation system, retrieval algorithm errors, and measurement time gaps (e.g., Anagnostou et al., 2001; Morris and Schwaller, 2011; Biswas and Chandrasekar, 2018; Warren et al., 2018). Following the procedure in Schwaller and Robert Morris (2011), we performed a sensitivity analysis of the four parameters described below on the variability of matchup samples. This analysis provides information on the sampling strategy that can reduce the aforementioned uncertainties and increase the reliability of CB estimates. The four parameters used in the sensitivity analysis are: (1) precipitation type; (2) partial beam filling ratio (PBFR); (3) vertical proximity to the mean bright band height defined as above, within, and below the bright band (BB); and (4) temporal gaps between DPR and GR measurements. All the parameters except for the temporal gaps can be estimated utilizing the variables provided by the DPR–GR geometry matched datasets. Table 1 summarizes the main variables in the DPR–GR matchups used in this study. For example, the mean BB layer (for each overpass) is defined as the DPR-identified mean BB height ± 0.75 km. This BB layer was used to categorize matchups' proximity to the BB as: (1) “above” if the bottom of a matching volume is above the BB layer; (2) “below” if the top of the volume is below the BB layer; and (3) “within,” otherwise. The PBFR indicates a proportion of the number of DPR and GR range bins (with the minimum reflectivity threshold of 15 dBZ for both of the DPR and a GR) within a matching volume. We note that the minimum detectable reflectivity of the DPR (Ku band) is reported as 12–13 dBZ (Skofronick-Jackson et al., 2018). The 15 dBZ threshold used above is a conservative value to avoid issues of variability of the DPR detection capability in different situations.

To quantitatively assess the effects of each parameter on the sample quality, we focused on the sample variability rather than the mean value of the reflectivity differences, which may vary over time. In cases where the temporal fluctuations of the mean values are severe, the standard deviation (σ) of all matchups for the entire sampling period may not be appropriate to represent the sample variability. Therefore, we also considered the weighted average of standard deviations (σ_w) in terms of sample size as a measure of the samples' dispersions.

$$\sigma_w = \sqrt{\frac{\sum \sigma_t^2 \times n_t}{\sum n_t}} \tag{6}$$

where σ_t^2 is the sample variance (dB^2) of GR – DPR at time t, n_t is the matchup sample size at time t, and σ_w is the weighted average of sample standard deviations (dB). The lower the variability of sample means are, the closer σ and σ_w are.

Among the four factors, the PBFR and the temporal separation are affected by characteristics of radar measurement mechanisms rather than precipitation system properties. Therefore, we first looked at the effects of these two factors and limited them with certain threshold values for further analyses, which consider various conditions defined by the permutations of the other factors.

Unlike other parameters, the accurate time separation between DPR and GR observations cannot be achieved based on the DPR–GR matchups because only the DPR's nearest approach time to GR and GR's scanning start time (and start time of each sweep) are available. We note that other earlier studies used the start time of either the entire volume scan (e.g., Wang and Wolff, 2009; Morris and Schwaller, 2011) or each elevation sweep (e.g. Schwaller and Robert Morris, 2011; Biswas and Chandrasekar, 2018; Warren et al., 2018) to estimate the time separation. However, the time period taken by a GR to complete a volume scan (i.e., 4–6 min in a precipitation mode of NEXRAD) was longer than the time it took for the DPR to scan the GR's coverage (i.e., ~30 s), thus degrading the accuracy of time gap measurements based on this approach. Clearly, the accuracy becomes worse when GR's volume scan starting time is used. For example, in the case of the DPR–GR matchups for the KDMX radar observed on June 26, 2015, the time differences between the DPR's nearest approach time and the KDMX volume scan starting time (DPR – GR) is 9 s, while the actual time separation for all matchup samples ranges from −221 to 18 s. We found that the actual time separation for part of the matchup samples could be > 13 min, even though the nearest approach time-based separation is < 5 min. Therefore, we recalculated more accurate time gaps for each matchup pair based on observation time records for each ray of the DPR and GRs. We retrieved the observation time information from the DPR Level 2 standard products and GR Level 2 volume scan data. Fig. 5 illustrates the effect of temporal separation between the DPR and KDMX on the agreement of two independent observations. Matchups of which time separations are < 30 s (Fig. 5(b)) clearly show less variability and stronger linear relationship (σ = 2.99, σ_w = 2.76, ρ = 0.93) than those without any filtering (σ = 3.96, σ_w = 3.83, ρ = 0.89 in Fig. 5(a)).

Fig. 6 illustrates the combined effects of the PBFR and the actual temporal measurement gaps on the statistics that reflect the differences between the DPR and the KDMX radar for the entire data period. As the PBFR increases and the measurement time difference decreases, the sample size decreases because of the stricter criteria applied (Fig. 6(a)). The changes in the variability can be found in Fig. 6(b) and (c), and the slightly different results between σ and σ_w can be attributed mainly to the variations of the mean differences over time. One observation is that the actual temporal gap is the dominant factor on sample variability in regions where the PBFR is > 50%. Consequently, using the 50% PBFR allows us to obtain more samples by adjusting the allowable time differences, while maintaining a similar variability level to that of the samples that are achievable only by applying a high PBFR threshold (i.e., 95% or 100%). Based on these results, we limited further analyses to the matchup samples of which PBFR values are ≥70% and the time differences are < 120 seconds, considering the sample size and variability (red points in Fig. 6). As a result, we obtained more reliable

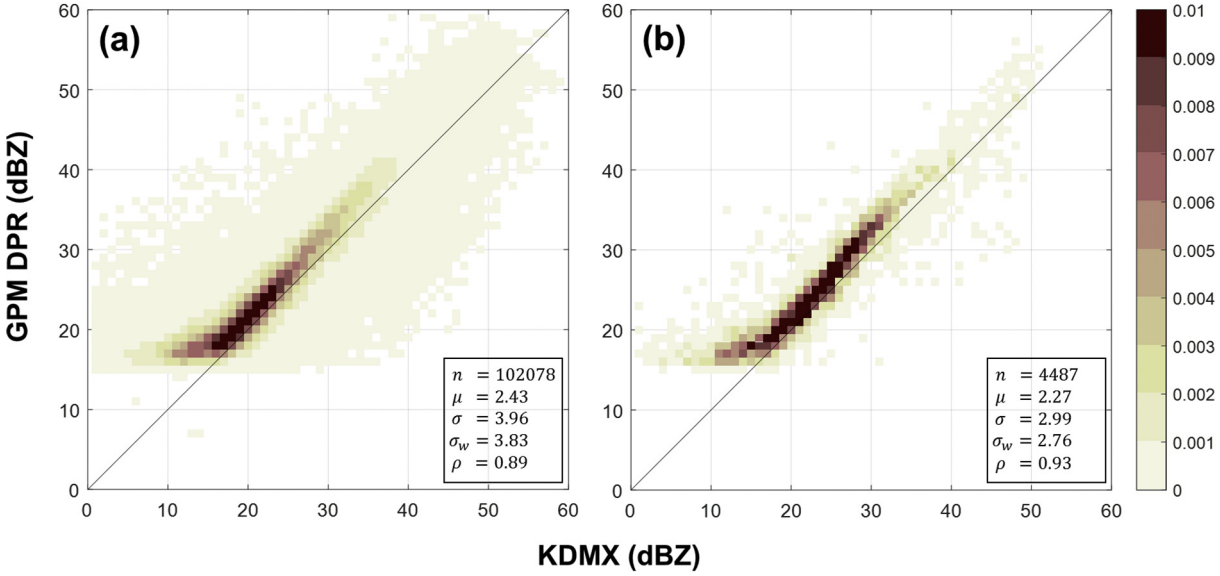

Fig. 5. Two-dimensional histogram plots for matchup samples over the entire period (a) without any filtering and (b) those of which actual measurement time offsets between the DPR and the KDMX radar are < 30 s. The colour scale represents frequency of the number of samples for each bin. The two-dimensional bin size is 1 dB.

matchup samples with a similar sample size, but less variability ($\sigma = 2.31$, $\sigma_w = 2.16$, $\rho = 0.96$) than the ones in which we used the 95% partial beam filling criterion as a limiting factor ($\sigma = 2.70$, $\sigma_w = 2.51$, $\rho = 0.93$).

With the PBFR and the temporal gap of measurements fixed at ≥70% and < 120 seconds, respectively, Table 2 illustrates the effects of different combinations of the other two factors, precipitation types (stratiform or convective), and proximity to the bright band (above, within, or below the BB) on sample variability. Table 2 shows that the variability of samples from the stratiform precipitation type is less than the convective type, regardless of the sampling regions related to the BB. Also, given precipitation types, samples above the BB show less variability. This could result from the uncertainties caused by the high spatial variability of convective precipitation system and the attenuation correction algorithm used in the DPR products (Morris and Schwaller, 2011). Among the six categories, taking samples classified as stratiform precipitation and above the BB leads to the smallest sample variability, which agrees with other relevant studies (e.g., Schwaller and Robert Morris, 2011). As a result, we used the DPR–GR matchups satisfying: (1) both of PBFR values of the DPR and GR of ≥70%; (2) the temporal gaps of measurements of < 120 seconds; (3) the stratiform precipitation type; and (4) above the BB.

We acknowledge that different frequencies between the DPR and GRs may also contribute to systematic reflectivity differences due to non-Rayleigh scattering effects (Anagnostou et al., 2001; Liao et al., 2001; Schwaller and Robert Morris, 2011; Wang and Wolff, 2009; Warren et al., 2018). These effects are a function of the size, phase, and density of the hydrometeors observed by radars and, in general, become more significant with the increase in reflectivity (or, increase in size and density of particles). To deal with this issue, we can use frequency conversion formulas (from Ku- to S-band or vice versa) derived from raindrop size distribution models (e.g., Liao and Meneghini, 2009; Cao et al., 2013) or regression-based bias estimation scheme dependent on reflectivity magnitude (e.g., Wang and Wolff, 2009). However, we assumed that this effect on the refined matchup samples is negligible and do not account for it in further analyses. This can be justified by the fact that the refined samples (above the BB and in stratiform precipitation) usually show low reflectivity of < 35 dBZ where the scattering effects are not significant (Anagnostou et al., 2001).

4.1.2. CB estimation

In this section, we present the CB estimates over time for the entire data period and selected GR sites as provided in Section 2. We eliminated overpasses for which the number of samples that satisfied the criteria explained above were < 25 from the analysis. As a result, we selected 10 out of 54, 7 out of 72, and 9 out of 61 overpasses for KDMX, KDVN, and KFSD NEXRAD radars, respectively.

Fig. 7 shows the CB estimates of three GRs against DPR measurements. Red squares represent the CB at each overpass, with a 95% confidence interval. The red solid lines represent the overall CB over the entire period for each radar; gray-shaded areas illustrate one standard deviation range of the CB estimates. Please note that we excluded the two extreme cases (July 26, 2015, and June 28, 2017) from the results for KFSD from the overall CB calculation. We will further discuss this shortly.

From Fig. 7, we observed that all three radars are negatively biased with overall biases of −1.89, −2.20, and −1.68 dB for KDMX, KDVN, and KFSD, respectively. It is interesting to note that their underestimation tendencies are consistent over time within ± 1 dB (the standard deviation values are 0.60 dB, 1.06 dB, and 0.42 dB for KDMX, KDVN, and KFSD, respectively). Only a few CB estimates deviate slightly from the one standard deviation interval. In fact, most of 95% confidence intervals of even those estimates overlap with one standard deviation range. This suggests that although all three GRs tend to systematically underestimate reflectivities when compared to the DPR measurements, the tendencies are consistent. We note that the increase in DPR reflectivity magnitude by the algorithm version change (refer to Section 2.1) indicates a certain level of bias may be included in DPR measurements even though it is assumed to be stable over time. Therefore, we may not conclude that the CB estimates are the same as the true calibration bias of the corresponding GRs. However, in the process of estimating the RCB between GRs, the overall bias involved in the DPR observations could be mitigated, so it may not be a problem in evaluating the GRs' relative biases. Note that the lack of DPR observations for rainy events is the main obstacle to analyzing temporally finer-scale variations of the CB. We will explain this issue in Section 4.2.

Let us examine closely the two extreme cases in Fig. 7 (c) to identify what factor caused such large differences. This is of interest not only because we use the reflectivity differences between the DPR and GRs as a measure of GRs' CB, but also because it could shed light on our speculation on KFSD's underestimation tendency when compared to

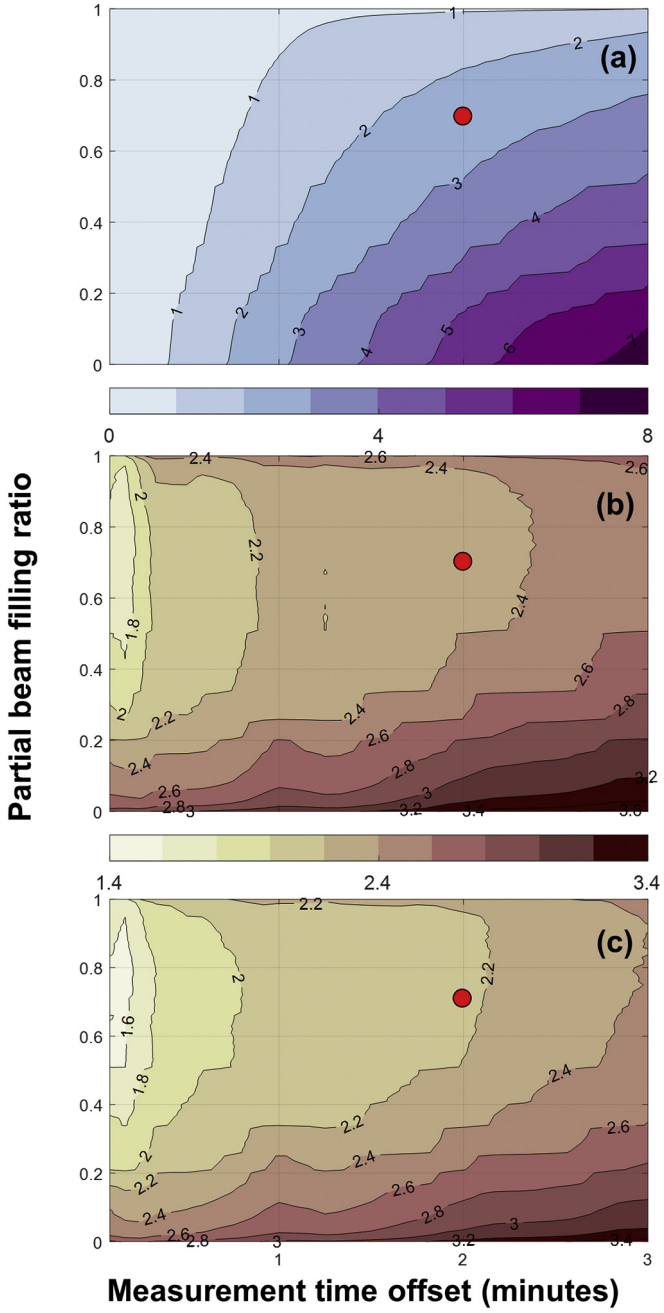

Fig. 6. Contour maps for (a) the number of matchup samples (sample size/10,000); (b) the ordinary sample standard deviation (dB); and (c) the weighted sample standard deviation (dB) of reflectivity differences between the DPR and the KDMX radar depending on changes in the criteria values of the PBFR and measurement time offsets (minutes). The red point in each contour map represent the location of corresponding sample statistics when the selected criteria (PBFR ≥70% and the time differences < 120 s) are applied to matchup samples. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

neighboring GRs (e.g., Fig. 1). We found that in both cases abrupt decreases in reflectivity magnitude occurred when a squall line sit on the top of the radar site (from Fig. 8(a) to (b) and from (d) to (e) for cases 1 and 2, respectively) and they were recovered after the squall lines passed over the radar (Fig. 8(c) and (f)). These temporal changes were quantitatively verified by area-weighted average reflectivity values estimated from the LEVEL II base scan reflectivity with > 10 dBZ within 100 km range from the radar (Fig. 8(g)). Both time series in Fig. 8(g) were centered on the KFSD's volume scan starting times. Although the

Table 2
Sample statistics of reflectivity differences depending on precipitation types and proximity to the bright band, given the PBFR of > 70% and measurement time offsets of < 2 min. This table excludes samples categorized as the "other" precipitation type by the 2ADPR algorithm.

Precipitation type	Proximity to the BB	σ	σ_w	ρ	N
Stratiform	Above	1.29	1.17	0.90	2533
	Within	1.95	1.86	0.94	10,384
	Below	1.75	1.58	0.96	3739
Convective	Above	2.35	2.27	0.96	1602
	Within	3.14	2.86	0.92	3241
	Below	3.67	3.39	0.88	1834

case 2 showed more complex pattern due to multiple strong cells passing over the radar, similar sudden decreases can be identified at time 0. These decreases are localized in time, which are unlikely to originate from natural precipitation system decay. Instead, we speculate that the observed decreases were caused by attenuation associated with wet radome effects. The existence of heavy rainfall over the KFSD site is supported by rainfall observations recorded by the Automated Surface Observing System (ASOS) rain gauge at Sioux Falls nearly 2 km away from the KFSD radar. This gauge shows that the maximum 1-minute rain rate was 107 mm/h (15 min rain rate of 73 mm/h) and 91 mm/h (40 mm/h) during the 15-minute period until the KFSD volume scans start on July 26, 2015, and June 28, 2017, respectively. Although wet radome or attenuation effects on S-band radars (i.e., NEXRAD) are conventionally ignored, two-way attenuation caused by wet radome effects can be severe (as much as 1 to 2 dB for heavy rainfall events [≈100 mm/h]) (Ryzhkov and Zrnic, 1995). Considering that the abruptly decreased status (1 to 2 dB) persists for < 1 h (e.g., Fig. 8(c) and (f)), the effects on rainfall accumulation would not be serious over a long period. However, the DPR–GR matchups contaminated by these attenuation-related effects can mislead the statistical inference on the CB (i.e., overestimation). Therefore, we excluded these two cases from our analysis. Please note that these examples suggested the need for a more detailed study on the GR's wet radome and attenuation issues because GR products are taken advantage of as a reference for the validation of the GPM DPR products (Schwaller and Robert Morris, 2011; Tapiador et al., 2012; Hou et al., 2014; Kim et al., 2014; Speirs et al., 2017; Kidd et al., 2018; Watters et al., 2018). However, these anomalous cases demonstrate the usefulness of the DPR measurements to monitor the abnormal behavior of GR observations.

4.2. Evaluation of RB

4.2.1. Detection of temporal change points

We performed multiple change-point analyses on the RB time series for the KDVN and KFSD radars against the KDMX to confirm their stationarity assumption over time. We selected RB estimates for this analysis instead of individual GRs' CB estimates primarily because of the abundancy of the RB estimates. We focused on the changes in the mean of the RB distribution over time.

In the case of the KDVN radar, the mean and standard deviation (SD) of RB over the entire period from 1 April 2014 to 30 September 2017 are 0.77 dB and 0.72 dB, respectively. For comparison, the corresponding RRCT values are 0.7 (mean) and 1.4 dB (SD), respectively. We attributed the smaller SD of our results to the temporal averaging. Fig. 9 (a) shows slight changes in the RB estimates over time, detected by the multiple change point analysis. As a result, we divided the RB time series into three statistically homogeneous periods in terms of mean by two statistically significant change points at 5% significance level on July 16, 2015, and July 7, 2017. We observed changes in the mean and SD values from 1.06 (period 1) to 0.53 (period 2) to 1.05 (period 3), and 0.71 to 0.61 to 0.89, respectively, as shown in Table 3. These changes suggest that at least one of the radars (KDMX and KDVN)

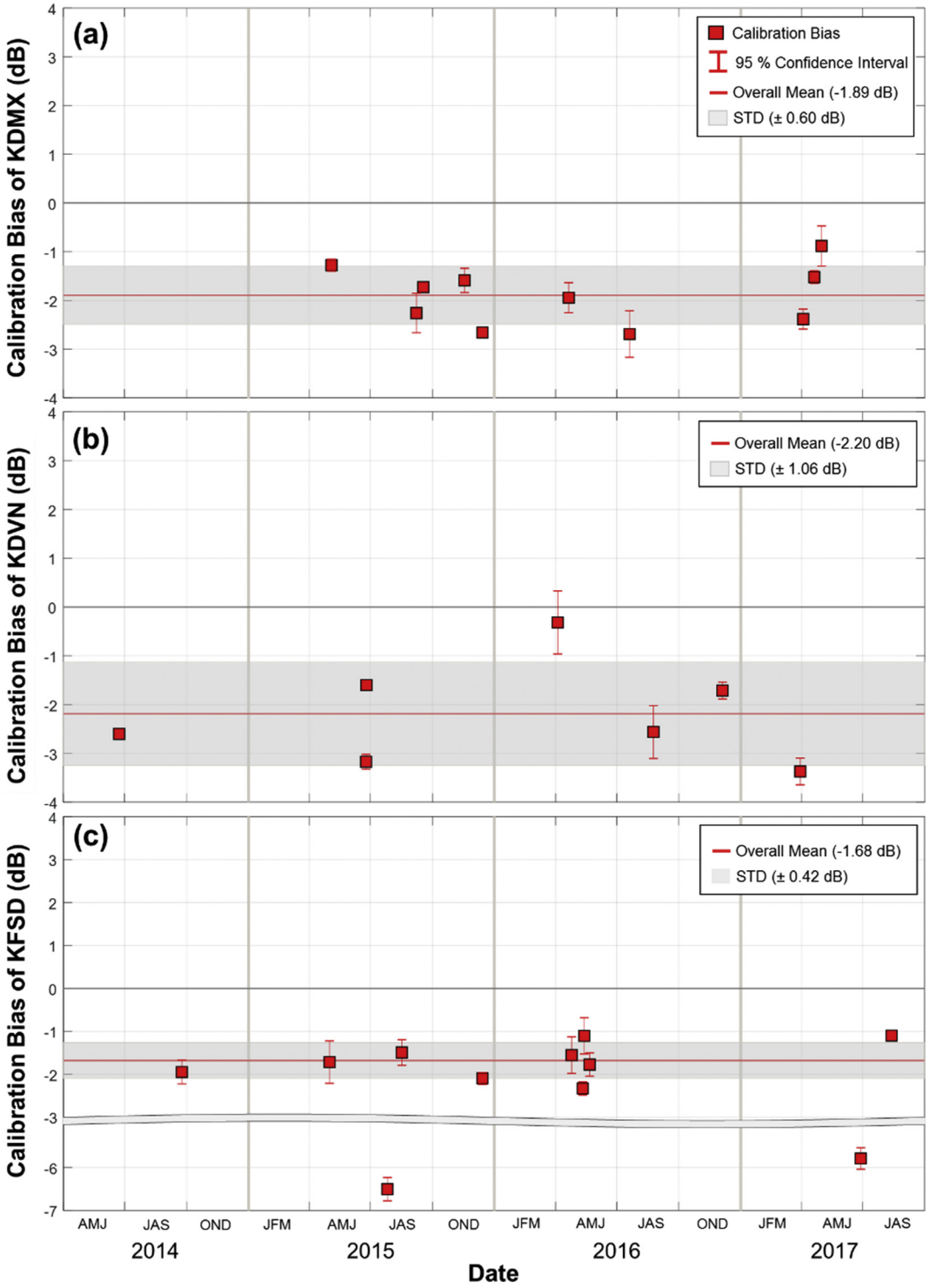

Fig. 7. Instantaneous calibration bias estimates of (a) KDMX, (b) KDVN, and (c) KFSD against the DPR measurements. The calibration bias is defined as the reflectivity difference between coincident GR and DPR (GR – DPR) in dBZ unit. Red squares represent the bias estimate, and the vertical bar shows the 95% confidence interval of the estimate. The mean CB over the entire period is shown as a red solid line, and its one standard deviation range is expressed by gray-shaded area. The CB values for events estimated by fewer than 25 matchup samples are excluded. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

may experience changes in the system performance because of gradual degradation of radar hardware over time or a calibration adjustment by routine maintenance procedures. Unfortunately, however, we could not verify this because we lacked detailed information on the radars' calibration maintenance.

The RB estimates for the KFSD radar (Fig. 9 (b)) show statistics similar (mean = 1.68 dB, SD = 1.15 dB) to those from RRCT (mean = 2.1 dB, SD = 1.6 dB). The most distinct feature of KFSD's RB estimates from the KDVN case is the seasonality, which can be detected by visual inspection as well as a multiple change point analysis. For an analysis on the seasonal difference, we categorized several periods into two categories: warm seasons (periods 1, 3, and 5); and cold seasons

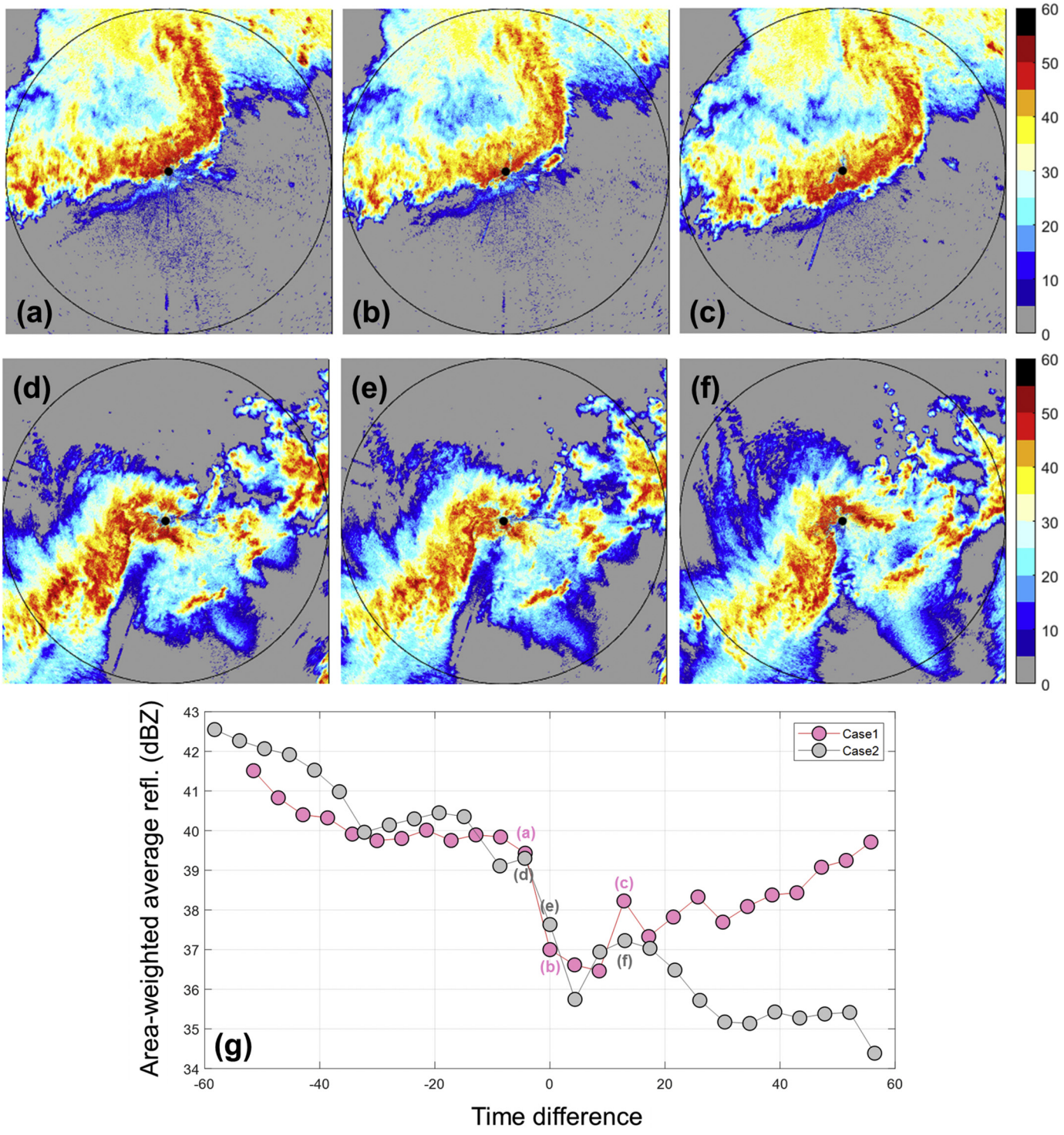

Fig. 8. Temporal changes in reflectivity observed by the KFSD radar for the two extreme cases. Base scan reflectivity maps are shown in top (case 1) and middle (case 2) rows: before ((a) and (d)) and after ((c) and (f) the squall lines passed over the radar site, and when the squall lines sit on the top of the radar site ((b) and (e)) at (a) 03:47:44 UTC, (b) 03:52:02 UTC, and (c) 04:04:54 UTC on July 26, 2015 and (d) 05:56:23 UTC, (e) 06:00:45 UTC, and (f) 06:13:46 UTC on June 28, 2017. Black circles represent 100 km range rings from the radar. Time series of the area-weighted average reflectivity above 10 dBZ within the 100 km range (g) are centered on the KFSD's volume scan starting times of case 1 and case 2.

(periods 2, 4, and 6), as shown in Fig. 9 (b) and Table 3. The mean RB values for each season show that they tend to be higher in cold seasons (2.10 dB) than in warm seasons (1.17 dB). Since the CB is likely due to internal system behavior, it is unreasonable to conjecture that the seasonal variability results from the CB. Because there is no seasonality in the comparison results between the KDMX and KDVN radars, we can infer that the KFSD radar suffers from other uncertainty sources affected by seasonal changes in environmental conditions (i.e., non-standard beam propagation, climatological effects, etc.). If such factors exist, they would be more influential on the mean structure than the variability considering small differences in the SD values (1.00 dB in warm season and 1.09 dB in cold season) in spite of the relatively large mean differences between the two seasons.

From the results of Section 4.1, we observed no clear evidence of the

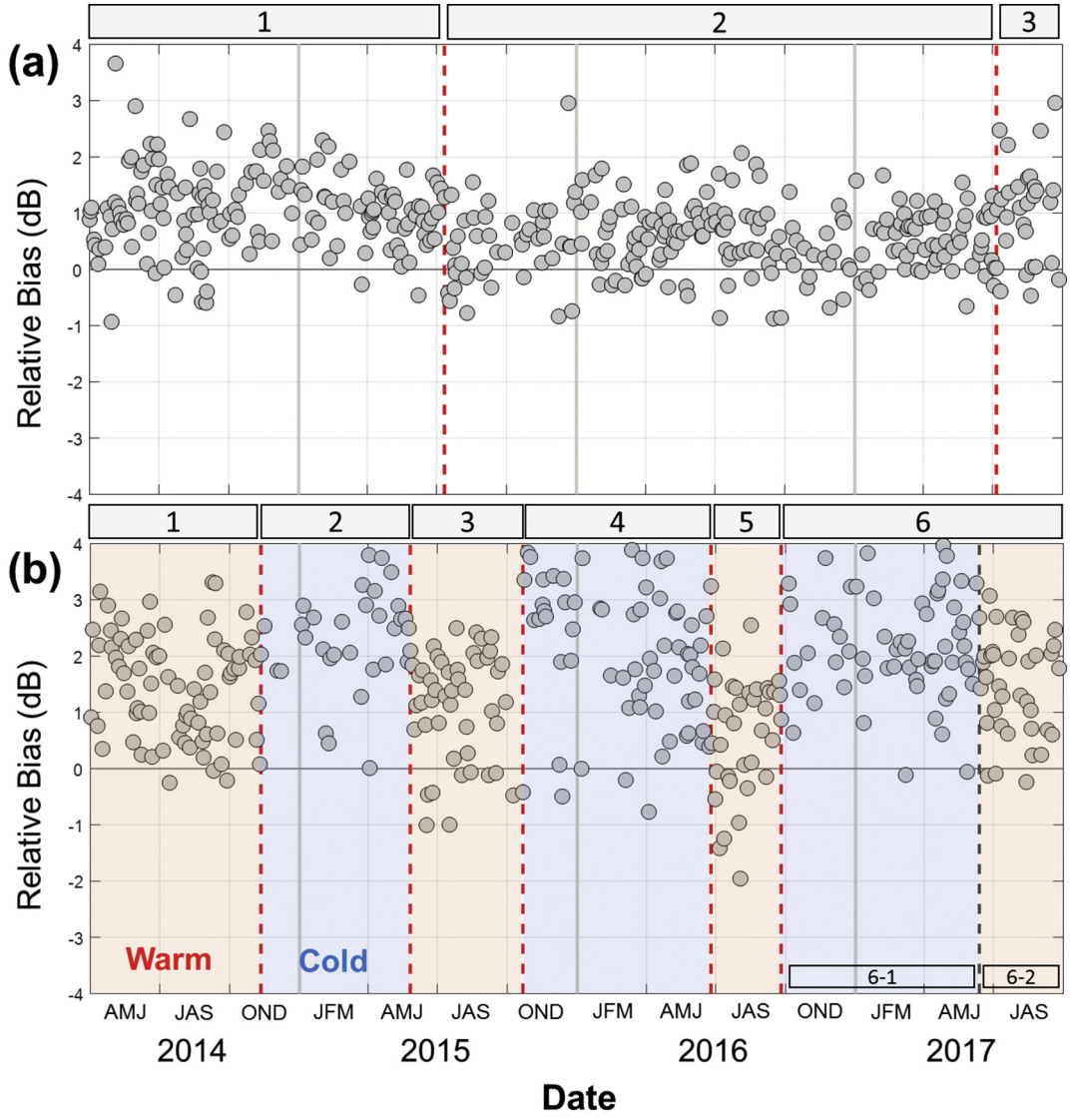

Fig. 9. The RB time series for the (a) KDVN and (b) KFSD radars against the reference radar (KDMX). The red vertical dotted lines show the distributional change point locations. Identified homogeneous periods are labeled by the numbers in light gray boxes at the top of each panel (also, see Table 3). The additional change point in period 6 of the KFSD's RB (b) detected by the second change point analysis (in Section 4.2.3) is marked by a dark gray dotted vertical line and labeled in bottom boxes in the panel. Light red and blue shaded area in (b) illustrate warm (May through October) and cold (November through April) seasons. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Distributional change point analysis results on the RB estimates of the KDVN and KFSD compared to the KDMX radar. Stationarity assumption on the RB holds for each period. Period 6* is more divided into two parts (6-1 and 6-2) by the second change point analysis on recollected RB estimates in terms of seasons.

GR 1	GR 2	Period index	Period		Relative bias (RB)		# of rain events	# of DPR overpasses for GR 1	# of DPR overpasses for GR 2	Season
			From	To	Mean (dB)	SD (dB)				
KDMX	KDVN	1	01-Apr-14	11-Jul-15	1.06	0.71	157	1	3	–
		2	16-Jul-15	06-Jul-17	0.53	0.61	224	9	4	–
		3	07-Jul-17	26-Sep-17	1.05	0.89	28	0	0	–
	KFSD	1	01-Apr-14	11-Nov-14	1.47	0.91	76	0	1	warm
		2	15-Nov-14	26-May-15	2.42	0.98	34	1	1	cold
		3	29-May-15	21-Oct-15	1.10	0.96	47	2	1	warm
		4	23-Oct-15	24-Jun-16	2.16	1.26	67	3	5	cold
		5	26-Jun-16	25-Sep-16	0.61	1.03	35	1	0	warm
		6*	04-Oct-16	25-Sep-17	1.94	0.98	98	3	1	–
		6-1	04-Oct-16	14-Jul-17	2.21	0.95	62	3	0	cold
		6-2	14-Jul-17	25-Sep-17	1.49	0.89	36	0	1	Warm

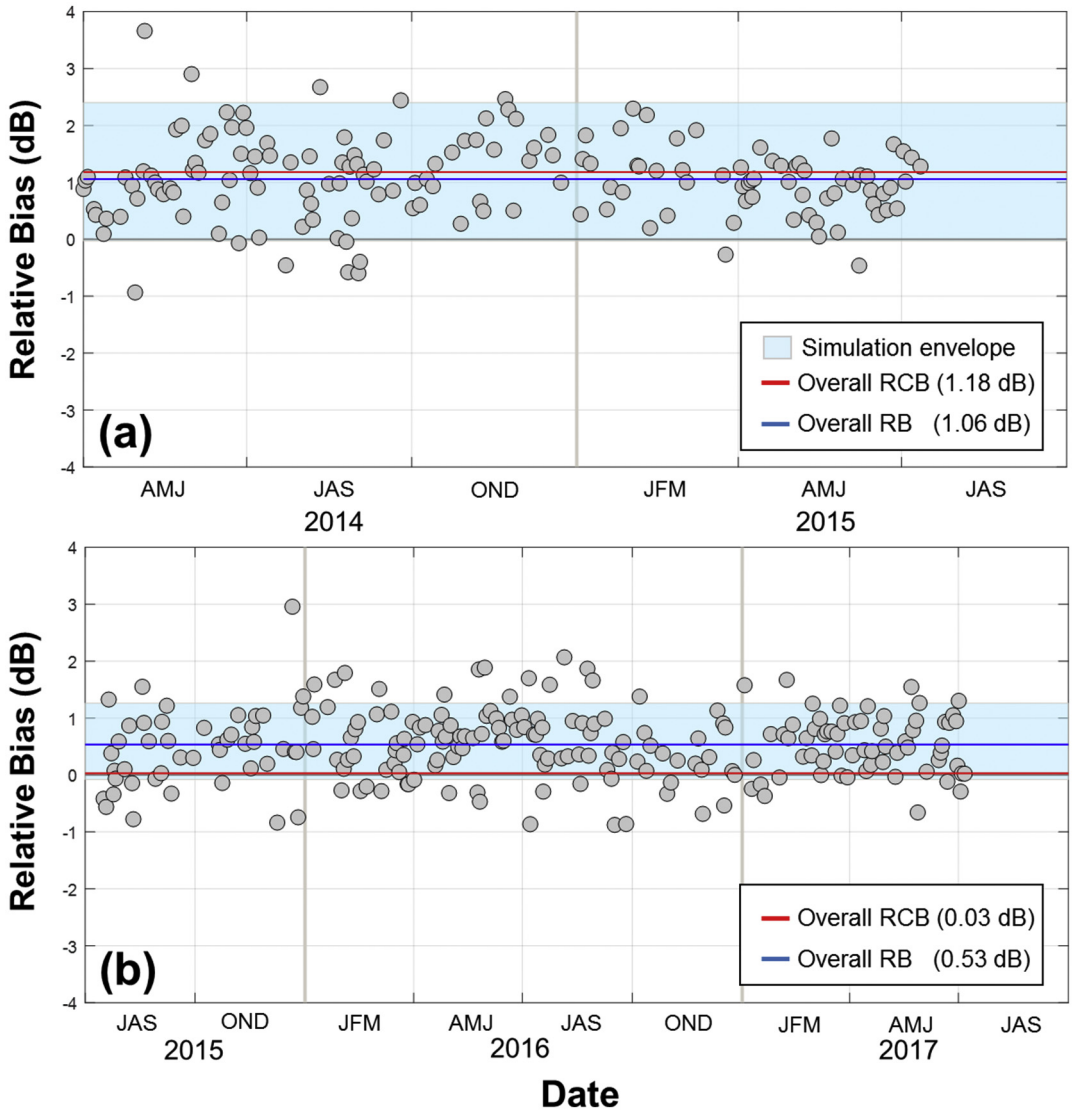

Fig. 10. The RB time series of the KDVN against the KDMX radar for (a) period 1 and (b) period 2 described in Table 3. Each point illustrates event-based RB estimates. Red and blue solid lines represent the overall RCB and RB over the periods, respectively. Light blue shaded area indicates the simulation envelope derived by the Monte Carlo simulations based on the RB. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

inconsistency in the CB of the three GRs against the DPR measurements over the entire time period. However, the results above show several change points in the RB time series, which implies that at least one among them underwent some changes. This suggests that infrequent DPR measurements have difficulty capturing progressive changes in the CB. Nevertheless, the DPR measurements could be useful to detect instantaneous systematic changes in GR observations (Fig. 8) and to evaluate the general performance of GRs from a long-term perspective.

4.2.2. RB between KDMX and KDVN

The Monte Carlo simulation results for periods 1 and 2 are illustrated in Fig. 10. Period 3 is excluded from the analysis since there is no DPR observation for both of the KDMX and KDVN radars. The mean RCB and RB estimates are very close, and the values are 1.18 and 1.06 dB for period 1 and 0.03 and 0.53 dB for period 2. The simulation envelopes, displayed as the blue shaded area, represent the range of possible mean RCB values inferred by RB estimates with the equivalent sample size of $\max[n_{KDMX-DPR}, n_{KDVN-DPR}]$ under the null hypothesis. With this in mind and because of sample size effects, we can observe that the range for period 1 with an equivalent sample size of 3 is clearly larger (−0.03 to 2.40 dB) than the range for period 2, with the equivalent sample size of 9 (−0.08 to 1.26 dB). Because the mean RCB values are within the simulation envelopes, the reflectivity differences observed at the equidistance zones are likely to be affected mainly by the calibration mismatch between the KDMX and KDVN radars. In this case, we can use either the RCB or RB estimates to mitigate the disagreements on the reflectivity of GRs.

4.2.3. RB between KDMX and KFSD

The seasonality of the RB estimates for the KFSD radar gives rise to the need to reorganize the original RB time series so we can secure more DPR observations while maintaining the validity of the stationarity assumption. From Section 4.2.1, we observed that the mean RB values categorized into the same season show the same pattern: high in the cold season and low in the warm season. We performed the multiple change point analysis for each season RB estimates. As a result, we detected only one change point for warm season at the starting point of period 5 (Table 3). This means that the mean of the RB distributions in period 1 and 3 are statistically the same (hereafter, defined as warm season RB estimates), but period 5 differs. This indicates that radar

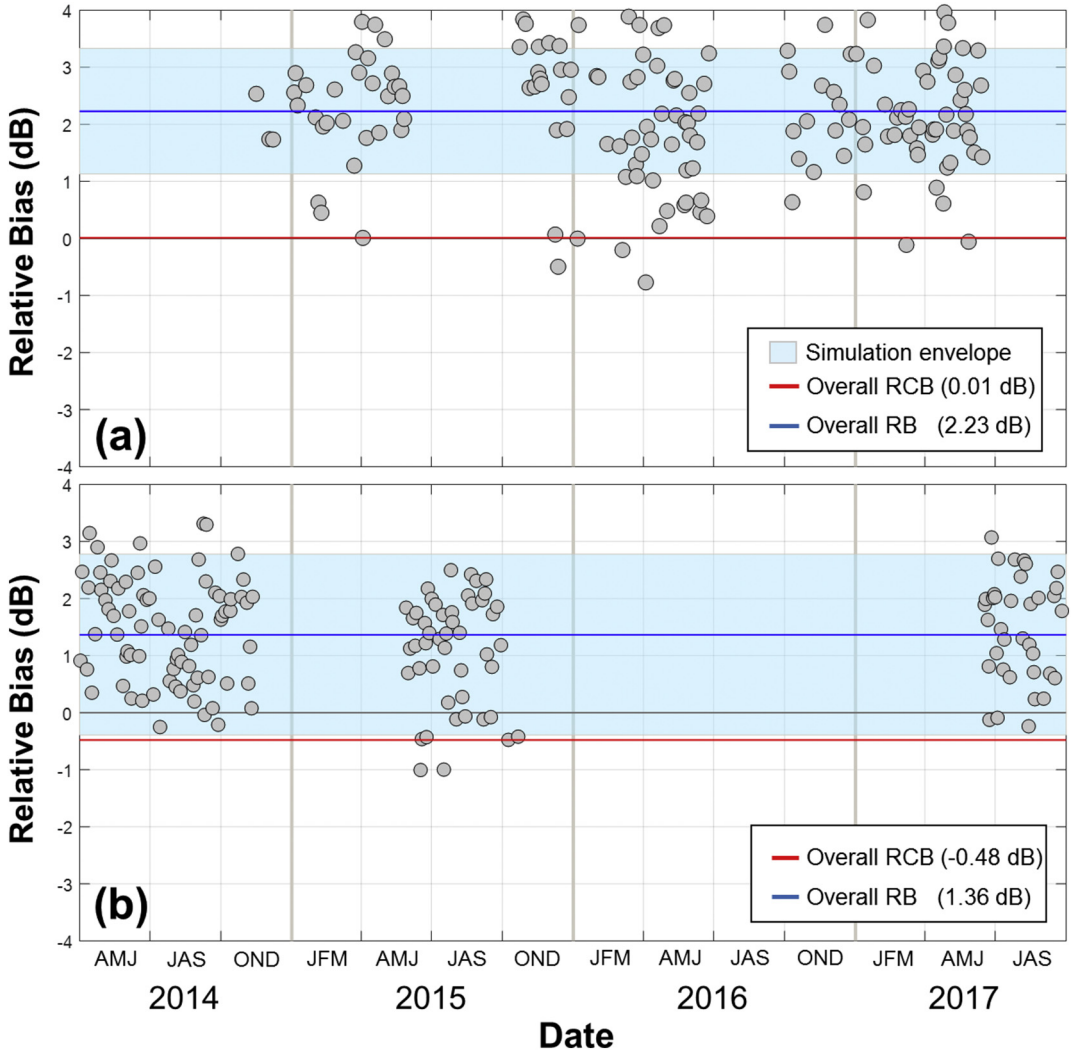

Fig. 11. Same as Fig. 10, but for the KFSD against the KDMX radar for (a) the cold season including periods 2, 4, and 6-1 and (b) the warm season with periods 1,3, and 6-2.

calibration for KDMX, KFSD, or both radars could be changed for period 5 as compared to other periods. However, it is impossible to further clarify the reason for the difference because only one DPR observation exists for the KDMX radar for period 5. In the cold season, one change point groups the RB time series into two parts: (1) period 2, 4, and the first 62 events in period 6 (that is, period 6-1); and (2) the last events in period 6. The former is hereafter defined as cold season RB estimates. We note that the change point in the middle of period 6 was not captured by the first change point analysis. We speculate that it would be caused mainly by the much smaller mean RB in period 5 (0.61 dB) than period 6 (1.94 dB). The large cross-variability between periods 5 and 6 made the inner-fluctuation of RB in period 6 seem to be relatively consistent so that it failed to detect the change point between periods 6-1 (2.21 dB) and 6-2 (1.49 dB). Due to statistical consistency of RB estimates in period 6-2 with those in periods 1 and 3, they were incorporated into the warm season RB estimates.

Fig. 11 clearly illustrates the difference in the mean RB values between the seasons (1.36 dB in warm and 2.23 dB in cold season). In contrast, the mean RCB shows relatively small difference between the seasons (−0.48 dB and 0.01 dB for the warm and cold season, respectively). The mean warm season RCB could include relatively larger uncertainties in representing the population mean because of the lack of DPR observations for both KDMX and KFSD radars ($n_{KDMX-DPR} = 2$, $n_{KFSD-DPR} = 3$). However, the results of small differences in the mean RCB estimates and large gaps in RB estimates between seasons suggest that other factors may affect the RB of the KFSD radar more than the calibration mismatch between the two radars does. As seen in Fig. 11 (b), the small sample size of DPR measurements gives rise to a wide simulation envelope (−0.39 dB to 2.78 dB) covering the variability of almost all individual RB estimates. This weakens the statistical inference on the null hypothesis. Therefore, we focused on the cold season data for further analysis.

Fig. 11 (a) displays the Monte Carlo simulation results for the cold season data. In this case, we estimated the mean cold season RB and RCB as 2.23 dB (blue line) and 0.01 dB (red line), respectively. The range of the simulation envelope (gray shaded area) is from 1.13 dB to 3.33 dB. We can observe the large difference of 2.22 dB between the RB and RCB, and the fact that the RCB is obviously outside the simulation envelope ($n_{eq} = 7$); this means that the radar calibration is clearly not the only factor in the RB and other factors are more significant.

Compared to the KDVN case (Section 4.2.2), the different characteristics of KFSD's RB, seasonality, and significant deviation from the mean RCB lead us to suspect the existence of other error sources rather than calibration bias affecting only the KFSD radar. In this context, we found that the topographic effect could be the most plausible explanation on this distinct feature of the KFSD radar. Unlike the KDMX and KDVN radars, the terrain surrounding the KFSD radar is sufficiently high enough above the radar site to give rise to partial beam blockages.

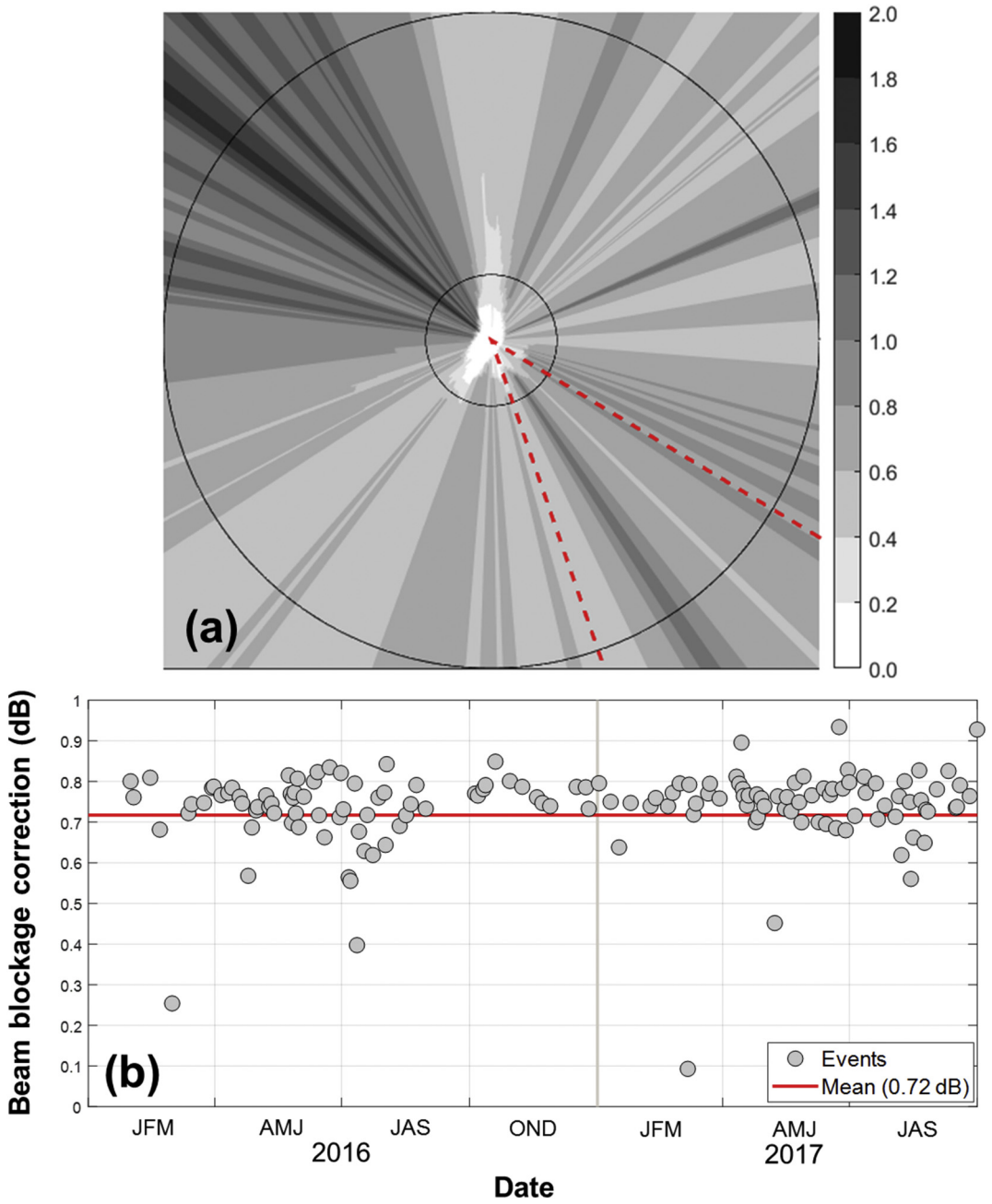

Fig. 12. (a) Power loss estimation (dB) for the KFSD radar caused by partial beam blockages based on the Gaussian beam pattern model with 0.5° elevation angle and 0.95° beam width. Inner and outer circles represent 10 km and 50 km range rings, respectively. Inner angles between two red-dotted lines shows the azimuth range of the ESV. (b) Reflectivity correction amount (dB) based on the beam blockage from January 2016 to September 2017. The red line represents the mean reflectivity correction amount (0.72 dB). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

To quantify the effects of the partial beam blockage on reflectivity products of the KFSD radar, we estimated these power losses based on the Gaussian beam pattern model with 0.5° elevation angle and 0.95° beam width. Note that, for the VN matchup data for NEXRAD radars, every volume-match sample has an associated value (GR_blockage variable in Table 1) indicating the mean fraction of GR beam blockage. The power distribution within the main lobe is not uniform, but instead depends on the distance of the blocked segments from the beam center. This cannot be accounted for by the mean fraction, so we considered the finer-scale interactions of each radar beam with the terrain, using 1/9th arc-second (approximately 3 m resolution) digital elevation models (DEMs) data from U.S. Geological Survey (U.S. Geological Survey, 2016). The power loss for each ray is integrated from −2 to 2°, both in azimuth and elevation angle directions, accounting for 97.7% of the total sensitivity of the beam pattern. Readers interested in the power loss simulation can consult Krajewski et al., 2006. The beam propagation, and thus the heights of radar range bins, are simulated under standard atmospheric conditions.

Fig. 12(a) shows the KFSD estimated power loss (dB) at 0.5° elevation angle in the native radar coordinates. We found that all radar beams within the azimuthal range of the ESV (from 109° to 149°, shown as red-dotted lines in Fig. 12(a)) are, in part, blocked by the terrain to a greater or lesser extent; the power loss ranges from 0.58 to 1.09 dB, with a mean of 0.78 dB.. The power loss is a deterministic function of the radar bins' range and azimuth in this simulation setup. Its effects on the RB are nonlinear, depending on the observed reflectivity magnitude

because the RB is estimated as a total reflectivity difference. To reflect this fact, we performed reflectivity corrections for the time period from September 2016 to September 2017 (Fig. 12(b)). As expected, the corrections reduced the mean RB by 0.72 dB on average (shown as the red line in Fig. 12(b)), which accounts for 50% (or more) of the mean RB values during warm seasons, but not the seasonal patterns. We speculated that the remaining parts and seasonality in the RB may originate mainly from more severe blockage caused by: (1) GRs' elevation angle of $< 0.5°$; and (2) non-standard beam propagations depending on environmental conditions (i.e., vertical gradients of refractivity). Please note that all blockages occur near the radar site (< 10 km), where radar sampling volume sizes are quite small. Fine deviations from the assumptions used in the simulation could have major effects on the intensity of the power loss. The KFSD radar's larger SD of RB values (0.93 for warm season and 1.09 for cold) compared to the KDVN radar (also see Table 3) could support this hypothesis. In addition, due to the longer range (≈ 160.5 km apart from both radars) of the ESV for the KDMX–KFSD comparison than the range (≈ 130.5 km) for the KDMX–KDVN, the elevations of GRs' base scan sampling volumes contributing to CAPPI products within the ESV for the KFSD case (3.36 km above sea level) are higher than the KDVN case (2.40 km). This is likely to exacerbate the blockage effects in combination with low level precipitation in cold seasons. In fact, it could increase the possibility of non-uniform beam filling and the occurrence of overshooting. Even though the RB was estimated based on GRs' vertically averaged reflectivity products targeted at the same altitude (i.e., CAPPI scan), observational limitations caused by these factors cannot be resolved. More accurate quantifications of each factor's effect on the RB could be another topic of research. Based on the results, however, it is obvious that the RB between adjacent radars with heterogeneous topographic or climatic conditions (i.e., KDMX vs. KFSD) could be dominated by other factors, including the partial beam blockage, rather than the radar calibration mismatch, even though they may not be substantially different. In this case, the bias correction based on the RCB cannot eliminate the RB. Also, systematic shifts of KFSD's reflectivity according to the RB estimate may unexpectedly intensify its bias magnitude because the partial beam blockage effects are dependent on the range and azimuth of radar bins.

We note that the partial beam blockage effects on the KFSD radar do not affect the CB and the RCB estimates in this study since the lowest elevation angle of the refined matchups (i.e., above the BB) is more than about 1.7°. For readers interested in more general DPR–GR comparisons, the blockage effect could be mitigated by: (1) excluding VN matchup samples with beam blockage above a specific threshold fraction; (2) incorporating quality information on the DPR or GR products into their comparisons (e.g., Crisologo et al., 2018); or (3) comparing the DPR instantaneous rainfall products with those derived by GR's dual-polarimetric algorithms less prone to the partial beam blockage effect.

5. Summary and conclusions

This study developed a methodology to compare reflectivity products from independent platforms (the GPM DPR and ground radars) to investigate reflectivity differences between adjacent GRs for the same targets. The inter-comparisons are performed as a way of assuring the agreement of sampling volumes by different radars in space and time. To compare different observational properties among the DPR and GRs, we developed a statistically integrated framework using inter-comparisons of all the platforms with a Monte Carlo simulation. To apply the proposed methodology, we collected the DPR–GR Normal Scan (NS) swath datasets (version 5) from the GVS VN and Iowa Flood Center's individual radar products for thee NEXRAD radars covering Iowa from April 2014 to September 2017. These radars include KDMX in Des Moines and KDVN in Davenport, Iowa; and KFSD in Sioux Falls, South Dakota. The following are the primary findings of our study:

1. We can reduce the observed variability in reflectivity between the DPR and GRs by constraining matchup samples on the following criteria: (1) PBFR $\geq 70\%$; (2) observation time offset ≤ 2 min; (3) precipitation type classified as stratiform; and (4) matchup elevation $\geq$ bright band (or melting layer) heights. The results consistently agreed with other studies, with the exception of the observation time gap effects. Our study achieved its unique results by computing the measurement time gaps based on the time differences of individual rays of the DPR and GRs. These proposed time gap measurements can accurately evaluate the variability caused by the time differences. The time gaps played an important role in determining the sample variability, given a PBFR $> 50\%$.
2. All three radars in the study (KDMX, KDVN, and KFSD) showed negative CB against the DPR observations for the entire time period, which means they systematically underestimated reflectivity compared to the DPR (on average, -1.89 dB, -2.20 dB, and -1.68 dB for KDMX, KDVN, and KFSD, respectively). However, the CB' standard deviation values for all the radars are around 1 dB or less (0.6 dB, 1.06 dB, and 0.42 dB for KDMX, KDVN, and KFSD), indicating that the underestimation tendencies are consistent over time. Because DPR observations can be biased at a certain level, the estimated calibration bias for each radar could deviate from the true calibration bias. This speculation is predicated on the increase (≈ 1.17 dB) in the DPR reflectivity caused by the DPR algorithm version change. However, the possible bias in the DPR reflectivity data could be canceled out in the process of estimating relative bias.
3. Two extreme cases prove the ability of the DPR to monitor system performance of GRs. The KFSD CB estimates on July 26, 2015, and June 28, 2017, showed exceptionally large decreases compared to others. These are caused by sudden decreases in the KFSD reflectivites when heavy rain fell on the radar site. We eliminated these cases from further analyses because we believed that the abnormal changes were likely caused by wet radome and the consequent attenuation effects rather than the CB.
4. Both the KDVN and KFSD RB time series, showed high temporal fluctuations when compared to the KDMX measurements. We detected several statistical change points in the mean using multiple change point analyses (two change points for the KDVN, and six for the KFSD). Periods divided by the change points showed distinct statistical behaviors, suggesting that the stationarity assumption for each period could be valid, but not for the entire analysis period. The infrequency of DPR overpasses made it impractical to identify such temporal variations by the DPR, giving rise to an obstacle to use of the DPR alone as an operational tool for the calibration bias corrections of the GRs.
5. Based on the simulation envelopes derived by Monte Carlo simulations, we tested the null hypothesis that the RB is caused by the RCB. In the case of the KDVN radar, the mean RB estimates were statistically the same as the mean RCB estimates. We concluded that the main source of the reflectivity differences between the KDMX and KDVN radars was likely the calibration mismatch. CB corrections based on either the RB or RCB (or CB for each GR) estimates can be expected to improve the reflectivity agreements between the two radars. In contrast, the RB of the KFSD radar against the KDMX showed statistically different characteristics from the RCB along with a conspicuous feature both in the mean and variability, i.e. seasonality. The discrepancy between the RB and RCB was greater in cold season. Because the radar calibration bias is mainly related to performance degradation of the radar system hardware and/or errors in the radar constant value at a given GR site, the seasonal trend is likely to be affected by other factors rather than the calibration mismatch. The most plausible explanation for the underestimation tendency of the KFSD radar was the topographic effects and consequent power losses caused by partial beam blockages. Under heterogeneous climatic or topographic conditions, factors other than calibration bias can dominate the reflectivity differences between

adjacent radars, requiring a different strategy to eliminate these differences.

Comparing our work to the study conducted by Warren et al. (2018), we can summarize the main differences as follows: (1) GRs' calibration bias changing point analysis based on RB estimates between adjacent GRs; and (2) diagnosis of the cause of the RB. Using engineering records on radar calibration maintenance, as Warren et al. did, helps identify actual changing points of GR calibration. However, the records are not readily obtainable for NEXRAD radars. Thus, our study focused on the temporal structures of RB estimates among GRs. In addition, our RB diagnosis (Section 4.2) can provide an appropriate remedy specific to the problem at hand, which may involve distinct error structures in space and time from the calibration bias. Given a situation where the RB is dominated by the calibration bias (i.e., adjacent GRs under similar environmental conditions as seen in our study), different methods provided similar results on the usefulness of the DPR observations to alleviate the bias, supporting the conclusions of both studies. Furthermore, our study focused on the validation of several issues underestimated by earlier studies, such as time gap effects on the noise of the DPR–GR matchups and negligible attenuation of S-band radars. We believe these issues are important to improve our understanding of miscalibration and other uncertainty sources in ground-based radar. The scarcity of the DPR overpasses makes it challenging to apply the approach presented in this study to operational radar QPE systems, especially when the RB is dominated by factors other than CB. Also, the variability of RB estimates makes it difficult to detect small changes in CB. One possible solution is to incorporate statistical characteristics of ground clutter echoes into the CB estimation procedure in this study (Silberstein et al., 2008; Wolff et al., 2015; Louf et al., 2019). Depending on the stability of the clutter echoes around a given GR site, this technique could detect even small changes in CB.

In the future, the wet-radome attenuation by S-band radars should be investigated as it is an underappreciated problem. GR's QPE algorithms based on polarimetric variables are well known as less prone to the attenuation effects. Still, more quantitative analyses on attenuation would help improve the accuracy of not only the DPR reflectivity-based algorithm but also single polarimetric GR applications still widely used over the world.

Acknowledgements

This study was supported by the Iowa Flood Center at the University of Iowa. We acknowledge and appreciate insightful discussions by Walt Petersen of NASA Marshall Space Flight Center.

References

Anagnostou, E.N., Morales, C.A., Dinku, T., 2001. The use of TRMM precipitation radar observations in determining ground radar calibration biases. J. Atmos. Ocean. Technol. 18, 616–628. https://doi.org/10.1175/1520-0426(2001)018<0616:TUOTPR>2.0.CO;2.

Atlas, D., 2002. Radar calibration some simple approaches. Bull. Am. Meteorol. Soc. 83, 1313–1316. https://doi.org/10.1175/1520-0477(2002)083<1313:RCSSA>2.3.CO;2.

Ayalew, T.B., Krajewski, W.F., Mantilla, R., 2014. Connecting the power-law scaling structure of peak-discharges to spatially variable rainfall and catchment physical properties. Adv. Water Resour. 71, 32–43. https://doi.org/10.1016/j.advwatres.2014.05.009.

Biswas, S., Chandrasekar, V., 2018. Cross-validation of observations between the GPM dual-frequency precipitation radar and ground based dual-polarization radars. Remote Sens. 10, 1773. https://doi.org/10.3390/rs10111773.

Bolen, S.M., Chandrasekar, V., 2003. Methodology for aligning and comparing space-borne radar and ground-based radar observations. J. Atmos. Ocean. Technol. 20, 647–659. https://doi.org/10.1175/1520-0426(2003)20<647:MFAACS>2.0.CO;2.

Brandes, E.A., Vivekanandan, J., Wilson, J.W., 1999. A comparison of radar reflectivity estimates of rainfall from collocated radars. J. Atmos. Ocean. Technol. 16, 1264–1272. https://doi.org/10.1175/1520-0426(1999)016<1264:ACORRE>2.0.CO;2.

Bringi, V.N., Rico-Ramirez, M.A., Thurai, M., 2011. Rainfall estimation with an operational polarimetric C-band radar in the United Kingdom: comparison with a gauge network and error analysis. J. Hydrometeorol. 12, 935–954. https://doi.org/10.1175/JHM-D-10-05013.1.

Cao, Q., Hong, Y., Qi, Y., Wen, Y., Zhang, J., Gourley, J.J., Liao, L., 2013. Empirical conversion of the vertical profile of reflectivity from Ku-band to S-band frequency. J. Geophys. Res. Atmos. 118, 1814–1825. https://doi.org/10.1002/jgrd.50138.

Crisologo, I., Warren, R.A., Mühlbauer, K., Heistermann, M., 2018. Enhancing the consistency of spaceborne and ground-based radar comparisons by using beam blockage fraction as a quality filter. Atmos. Meas. Tech. 11, 5223–5236. https://doi.org/10.5194/amt-11-5223-2018.

Cunha, L.K., Smith, J.A., Krajewski, W.F., Baeck, M.L., Seo, B.-C., 2015. NEXRAD NWS polarimetric precipitation product evaluation for IFloodS. J. Hydrometeorol. 16, 1676–1699. https://doi.org/10.1175/JHM-D-14-0148.1.

Frech, M., Hagen, M., Mammen, T., 2017. Monitoring the absolute calibration of a polarimetric weather radar. J. Atmos. Ocean. Technol. 34, 599–615. https://doi.org/10.1175/JTECH-D-16-0076.1.

Gabella, M., Morin, E., Notarpietro, R., Michaelides, S., 2013. Winter precipitation fields in the Southeastern Mediterranean area as seen by the Ku-band spaceborne weather radar and two C-band ground-based radars. Atmos. Res. 119, 120–130. https://doi.org/10.1016/j.atmosres.2011.06.001.

Hou, A.Y., Kakar, R.K., Neeck, S., Azarbarzin, A.A., Kummerow, C.D., Kojima, M., Oki, R., Nakamura, K., Iguchi, T., 2014. The global precipitation measurement mission. Bull. Am. Meteorol. Soc. 95, 701–722. https://doi.org/10.1175/BAMS-D-13-00164.1.

Kidd, C., Tan, J., Kirstetter, P.E., Petersen, W.A., 2018. Validation of the version 05 level 2 precipitation products from the GPM core observatory and constellation satellite sensors. Q. J. R. Meteorol. Soc. 144, 313–328. https://doi.org/10.1002/qj.3175.

Kim, J.H., Ou, M.L., Park, J.D., Morris, K.R., Schwaller, M.R., Wolff, D.B., 2014. Global precipitation measurement (GPM) ground validation (GV) prototype in the Korean Peninsula. J. Atmos. Ocean. Technol. 31, 1902–1921. https://doi.org/10.1175/JTECH-D-13-00193.1.

Krajewski, W.F., Ntelekos, A.A., Goska, R., 2006. A GIS-based methodology for the assessment of weather radar beam blockage in mountainous regions: two examples from the US NEXRAD network. Comput. Geosci. 32, 283–302. https://doi.org/10.1016/j.cageo.2005.06.024.

Krajewski, W.F., Ceynar, D., Demir, I., Goska, R., Kruger, A., Langel, C., Mantillllla, R., Niemeier, J., Quintero, F., Seo, B.C., Smallll, S.J., Weber, L.J., Young, N.C., 2017. Real-time flood forecasting and information system for the state of Iowa. Bull. Am. Meteorol. Soc. 98, 539–554. https://doi.org/10.1175/BAMS-D-15-00243.1.

Liao, L., Meneghini, R., 2009. Validation of TRMM precipitation radar through comparison of its multiyear measurements with ground-based radar. J. Appl. Meteorol. Climatol. 48, 804–817. https://doi.org/10.1175/2008JAMC1974.1.

Liao, L., Meneghini, R., Iguchi, T., 2001. Comparisons of rain rate and reflectivity factor derived from the TRMM precipitation radar and the WSR-88D over the Melbourne, Florida, site. J. Atmos. Ocean. Technol. 18, 1959–1974. https://doi.org/10.1175/1520-0426(2001)018<1959:CORRAR>2.0.CO;2.

Louf, V., Protat, A., Warren, R.A., Collis, S.M., Wolff, D.B., Raunyiar, S., Jakob, C., Petersen, W.A., 2019. An integrated approach to weather radar calibration and monitoring using ground clutter and satellite comparisons. J. Atmos. Ocean. Technol. 36, 17–39. https://doi.org/10.1175/JTECH-D-18-0007.1.

Matteson, D.S., James, N.A., 2014. A nonparametric approach for multiple change point analysis of multivariate data. J. Am. Stat. Assoc. 109, 334–345. https://doi.org/10.1080/01621459.2013.849605.

Mikosch, T., Kallenberg, O., 1998. Foundations of modern probability. J. Am. Stat. Assoc. 93, 1243. https://doi.org/10.2307/2669881.

Morris, K.R., Schwaller, M.R., 2011. Sensitivity of spaceborne and ground radar comparison results to data analysis methods and constraints. In: 35th Conf. On Radar Meteorology. Amer. Meteor. Soc, Pittsburgh, PA, pp. 68.

National Aeronautics and Space Administration, 2015. Global Precipitation Mission (GPM) Ground Validation System Validation Network Data Product User's Guide. Goddard Space Flight Center, Greenbelt, Maryland Retrieved from. https://pmm.nasa.gov/sites/default/files/document_files/Val_Network_Users_Guide_Vol_2_Nov2015.pdf, Accessed date: 18 January 2018.

Quintero, F., Krajewski, W.F., Mantilla, R., Small, S., Seo, B.-C., 2016. A spatial–dynamical framework for evaluation of satellite rainfall products for flood prediction. J. Hydrometeorol. 17, 2137–2154. https://doi.org/10.1175/JHM-D-15-0195.1.

Ryzhkov, A., Zrnic, D.S., 1995. Precipitation and attenuation measurements at a 10-cm wavelength. J. Appl. Meteorol. 34, 2121–2134. https://doi.org/10.1175/1520-0450(1995)034<2120:PAAMAA>2.0.CO;2.

Schwaller, M.R., Robert Morris, K., 2011. A ground validation network for the global precipitation measurement mission. J. Atmos. Ocean. Technol. 28, 301–319. https://doi.org/10.1175/2010JTECHA1403.1.

Seo, B., Krajewski, W.F., Kruger, A., Domaszczynski, P., Smith, J.A., Steiner, M., 2011. Radar-rainfall estimation algorithms of hydro-NEXRAD. J. Hydroinf. 13, 277–291. https://doi.org/10.2166/hydro.2010.003.

Seo, B.C., Cunha, L.K., Krajewski, W.F., 2013. Uncertainty in radar-rainfall composite and its impact on hydrologic prediction for the eastern Iowa flood of 2008. Water Resour. Res. 49, 2747–2764. https://doi.org/10.1002/wrcr.20244.

Seo, B.-C., Krajewski, W.F., Smith, J.A., 2014. Four-dimensional reflectivity data comparison between two ground-based radars: methodology and statistical analysis. Hydrol. Sci. J. 59, 1320–1334. https://doi.org/10.1080/02626667.2013.839872.

Seo, B.C., Krajewski, W.F., Mishra, K.V., 2015. Using the new dual-polarimetric capability of WSR-88D to eliminate anomalous propagation and wind turbine effects in radar-rainfall. Atmos. Res. 153, 296–309. https://doi.org/10.1016/j.atmosres.2014.09.004.

Silberstein, D.S., Wolff, D.B., Marks, D.A., Atlas, D., Pippitt, J.L., 2008. Ground clutter as a monitor of radar stability at Kwajalein, RMI. J. Atmos. Ocean. Technol. 25, 2037–2045. https://doi.org/10.1175/2008JTECHA1063.1.

Skofronick-Jackson, G., Berg, W., Kdd, C., Kirschbaum, D.B., Petersen, W.A., Juffman, G.J., Takayabu, Y.N., 2018. Global precipitation measurement (GPM): Unified precipitation estimation from space. In: Andronache, C. (Ed.), Remote Sensing of Clouds and Precipitation. Springer International Publishing, Cham, Switzerland, 978-3-319-72583-3, pp. 175–193.

Smith, J.A., Seo, D.J., Baeck, M.L., Hudlow, M.D., 1996. An intercomparison study of NEXRAD precipitation estimates. Water Resour. Res. 32, 2035–2045. https://doi.org/10.1029/96WR00270.

Speirs, P., Gabella, M., Berne, A., 2017. A comparison between the GPM dual-frequency precipitation radar and ground-based radar precipitation rate estimates in the Swiss Alps and Plateau. J. Hydrometeorol. 18, 1247–1269. https://doi.org/10.1175/JHM-D-16-0085.1.

Tapiador, F.J., Turk, F.J., Petersen, W., Hou, A.Y., García-Ortega, E., Machado, L.A.T., Angelis, C.F., Salio, P., Kidd, C., Huffman, G.J., de Castro, M., 2012. Global precipitation measurement: methods, datasets and applications. Atmos. Res. 104–105, 70–97. https://doi.org/10.1016/j.atmosres.2011.10.021.

Thurai, M., Mishra, K.V., Bringi, V.N., Krajewski, W.F., 2017. Initial results of a new composite-weighted algorithm for dual-polarized X-band rainfall estimation. J. Hydrometeorol. 18, 1081–1100. https://doi.org/10.1175/JHM-D-16-0196.1.

U.S. Geological Survey, 2016. USGS National Elevation Dataset (NED) 1/9 arc-second Downloadable Data Collection from the National Map 3D Elevation Program (3DEP) - National Geospatial Data Asset (NGDA) National Elevation Data Set (NED). U.S. Geological Survey.

Villarini, G., Krajewski, W.F., 2010. Review of the different sources of uncertainty in single polarization radar-based estimates of rainfall. Surv. Geophys. 31, 107–129. https://doi.org/10.1007/s10712-009-9079-x.

Wang, J., Wolff, D.B., 2009. Comparisons of reflectivities from the TRMM precipitation radar and ground-based radars. J. Atmos. Ocean. Technol. 26, 857–875. https://doi.org/10.1175/2008JTECHA1175.1.

Warren, R.A., Protat, A., Siems, S.T., Ramsay, H.A., Louf, V., Manton, M.J., Kane, T.A., 2018. Calibrating ground-based radars against TRMM and GPM. J. Atmos. Ocean. Technol. 35, 323–346. https://doi.org/10.1175/JTECH-D-17-0128.1.

Watters, D., Battaglia, A., Mroz, K., Tridon, F., 2018. Validation of the GPM version-5 surface rainfall products over Great Britain and Ireland. J. Hydrometeorol. https://doi.org/10.1175/JHM-D-18-0051.1.

Wolff, D.B., Marks, D.A., Petersen, W.A., 2015. General application of the relative calibration adjustment (RCA) technique for monitoring and correcting radar reflectivity calibration. J. Atmos. Ocean. Technol. 32, 496–506. https://doi.org/10.1175/JTECH-D-13-00185.1.

Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., Grams, H., Wang, Y., Cockcks, S., Martinaitis, S., Arthur, A., Cooper, K., Brogden, J., Kitzmilllller, D., 2016. Multi-radar multi-sensor (MRMS) quantitative precipitation estimation: initial operating capabilities. Bull. Am. Meteorol. Soc. 97, 621–638. https://doi.org/10.1175/BAMS-D-14-00174.1.

Zhong, L., Yang, R., Wen, Y., Chen, L., Gou, Y., Li, R., Zhou, Q., Hong, Y., 2017. Cross-evaluation of reflectivity from the space-borne precipitation radar and multi-type ground-based weather radar network in China. Atmos. Res. 196, 200–210. https://doi.org/10.1016/j.atmosres.2017.06.016.